APPENDIX C

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
23 April 2009 (23.04.2009)

PCT

(10) International Publication Number
WO 2009/049350 A1

(51) International Patent Classification:
C12N 15/11 (2006.01)

(21) International Application Number:
PCT/AU2008/001463

(22) International Filing Date: 2 October 2008 (02.10.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/980,145 15 October 2007 (15.10.2007) US

(71) Applicant *(for all designated States except US)*: **THE UNIVERSITY OF QUEENSLAND** [AU/AU]; St Lucia, QLD 4067 (AU).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **FRAZER, Ian Hector** [AU/AU]; 14 Jerdanefield Street, St Lucia, QLD 4067 (AU).

(74) Agents: **ARGAET, Victor P.** et al.; Davies Collison Cave, Level 3, 303 Coronation Drive, Milton, Queensland 4064 (AU).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

WO 2009/049350 A1

(54) Title: EXPRESSION SYSTEM FOR MODULATING AN IMMUNE RESPONSE

(57) Abstract: The present invention discloses methods and compositions for modulating the quality of an immune response to a target antigen in a mammal, which response results from the expression of a polynucleotide that encodes at least a portion of the target antigen, wherein the quality is modulated by replacing at least one codon of the polynucleotide with a synonymous codon that has a higher or lower preference of usage by the mammal to confer the immune response than the codon it replaces.

"EXPRESSION SYSTEM FOR MODULATING AN IMMUNE RESPONSE"

FIELD OF THE INVENTION

[0001] The present invention relates generally to gene expression. More particularly, the present invention relates to methods for modulating the quality of an immune response to a target antigen in a mammal, which response results from the expression of a polynucleotide that encodes at least a portion of the target antigen, wherein the quality is modulated by replacing at least one codon of the polynucleotide with a synonymous codon that has a higher or lower preference of usage by the mammal to confer the immune response than the codon it replaces. Even more particularly, the present invention relates to the use of a protein-encoding polynucleotide whose codon composition has been modified for modulating the quality of an immune response to an antigen in a mammal.

BACKGROUND OF THE INVENTION

[0002] The expression of foreign heterologous genes in transformed cells is now commonplace. A large number of mammalian genes, including, for example, murine and human genes, have been successfully expressed in various host cells, including bacterial, yeast, insect, plant and mammalian host cells. Nevertheless, despite the burgeoning knowledge of expression systems and recombinant DNA technology, significant obstacles remain when one attempts to express a foreign or synthetic gene in a selected host cell. For example, translation of a synthetic gene, even when coupled with a strong promoter, often proceeds much more slowly than would be expected. The same is frequently true of exogenous genes that are foreign to the host cell. This lower than expected translation efficiency is often due to the protein coding regions of the gene having a codon usage pattern that does not resemble those of highly expressed genes in the host cell. It is known in this regard that codon utilization is highly biased and varies considerably in different organisms and that biases in codon usage can alter peptide elongation rates. It is also known that codon usage patterns are related to the relative abundance of tRNA isoacceptors, and that genes encoding proteins of high versus low abundance show differences in their codon preferences.

[0003] The implications of codon preference phenomena on gene expression are manifest in that these phenomena can affect the translational efficiency of messenger RNA (mRNA). It is widely known in this regard that translation of "rare codons", for which the

corresponding iso-tRNA is in low abundance relative to other iso-tRNAs, may cause a ribosome to pause during translation which can lead to a failure to complete a nascent polypeptide chain and an uncoupling of transcription and translation. Thus, the expression of an exogenous gene may be impeded severely if a particular host cell of an organism or the organism itself has a low abundance of iso-tRNAs corresponding to one or more codons of the exogenous gene. Accordingly, a major aim of investigators in this field is to first ascertain the codon preference for particular cells in which an exogenous gene is to be expressed, and to subsequently alter the codon composition of that gene for optimized expression in those cells.

[0004] Codon-optimization techniques are known for improving the translational kinetics of translationally inefficient protein coding regions. Traditionally, these techniques have been based on the replacement of codons that are rarely or infrequently used in the host cell with those that are host-preferred. Codon frequencies can be derived from literature sources for the highly expressed genes of many organisms (see, for example, Nakamura *et al.*, 1996, Nucleic Acids Res 24: 214-215). These frequencies are generally expressed on an 'organism-wide average basis' as the percentage of occasions that a synonymous codon is used to encode a corresponding amino acid across a collection of protein-encoding genes of that organism, which are preferably highly expressed.

[0005] Typically, codons are classified as: (a) "common" codons (or "preferred" codons) if their frequency of usage is above about 4/3 × the frequency of usage that would be expected in the absence of any bias in codon usage; (b) "rare" codons (or "non-preferred" codons) if their frequency of usage is below about 2/3 × the frequency of usage that would be expected in the absence of any bias in codon usage; and (c) "intermediate" codons (or "less preferred" codons) if their frequency of usage is in-between the frequency of usage of "common" codons and of "rare" codons. Since an amino acid can be encoded by 2, 3, 4 or 6 codons, the frequency of usage of any selected codon, which would be expected in the absence of any bias in codon usage, will be dependent upon the number of synonymous codons which code for the same amino acid as the selected codon. Accordingly, for a particular amino acid, the frequency thresholds for classifying codons in the "common", "intermediate" and "rare" categories will be dependent upon the number of synonymous codons for that amino acid. Consequently, for amino acids having 6 choices of synonymous codon, the frequency of codon usage that would be expected in the absence of any bias in codon usage is 16% and thus the "common", "intermediate" and "rare" codons are defined as

those codons that have a frequency of usage above 20%, between 10 and 20% and below 10%, respectively. For amino acids having 4 choices of synonymous codon, the frequency of codon usage that would be expected in the absence of codon usage bias is 25% and thus the "common", "intermediate" and "rare" codons are defined as those codons that have a frequency of usage above 33%, between 16 and 33% and below 16%, respectively. For isoleucine, which is the only amino acid having 3 choices of synonymous codon, the frequency of codon usage that would be expected in the absence of any bias in codon usage is 33% and thus the "common", "intermediate" and "rare" codons for isoleucine are defined as those codons that have a frequency of usage above 45%, between 20 and 45% and below 20%, respectively. For amino acids having 2 choices of synonymous codon, the frequency of codon usage that would be expected in the absence of codon usage bias is 50% and thus the "common", "intermediate" and "rare" codons are defined as those codons that have a frequency of usage above 60%, between 30 and 60% and below 30%, respectively. Thus, the categorization of codons into the "common", "intermediate" and "rare" classes (or "preferred", "less preferred" or "non preferred", respectively) has been based conventionally on a compilation of codon usage for an organism in general (*e.g.*, 'human-wide') or for a class of organisms in general (*e.g.*, 'mammal-wide'). For example, reference may be made to Seed (see U.S. Patent Serial Nos 5,786,464 and 5,795,737) who discloses preferred, less preferred and non-preferred codons for mammalian cells in general. However, the present inventor revealed in WO 99/02694 and in WO 00/42190 that there are substantial differences in the relative abundance of particular iso-tRNAs in different cells or tissues of a single multicellular organism (*e.g.*, a mammal or a plant) and that this plays a pivotal role in protein translation from a coding sequence with a given codon usage or composition.

[0006] Thus, in contrast to the art-recognized presumption that different cells of a multicellular organism have the same bias in codon usage, it was revealed for the first time that one cell type of a multicellular organism uses codons in a manner distinct from another cell type of the same organism. In other words, it was discovered that different cells of an organism can exhibit different translational efficiencies for the same codon and that it was not possible to predict which codons would be preferred, less preferred or non preferred in a selected cell type. Accordingly, it was proposed that differences in codon translational efficiency between cell types could be exploited, together with codon composition of a gene, to regulate the production of a protein in, or to direct that production to, a chosen cell type.

[0007] Therefore, in order to optimize the expression of a protein-encoding polynucleotide in a particular cell type, WO 99/02694 and in WO 00/42190 teach that it is necessary to first determine the translational efficiency for each codon in that cell type, rather than to rely on codon frequencies calculated on an organism-wide average basis, and then to codon modify the polynucleotide based on that determination.

[0008] The present inventor further disclosed in WO 2004/042059 a strategy for enhancing or reducing the quality of a selected phenotype that is displayed, or proposed to be displayed, by an organism of interest. The strategy involves codon modification of a polynucleotide that encodes a phenotype-associated polypeptide that either by itself, or in association with other molecules, in the organism of interest imparts or confers the selected phenotype upon the organism. Unlike previous methods, however, this strategy does not rely on data that provide a ranking of synonymous codons according to their preference of usage in an organism or class of organisms. Nor does it rely on data that provide a ranking of synonymous codons according to their translational efficiencies in one or more cells of the organism or class of organisms. Instead, it relies on ranking individual synonymous codons that code for an amino acid in the phenotype-associated polypeptide according to their preference of usage by the organism or class of organisms, or by a part thereof, for producing the selected phenotype.

SUMMARY OF THE INVENTION

[0009] The present invention is predicated in part on the experimental determination of a ranking of individual synonymous codons according to their preference for producing an immune response, including a humoral immune response, to an antigen in a mammal. Significantly, this ranking is not coterminous with a ranking of codon frequency values derivable from an analysis of the frequency with which codons are used to encode their corresponding amino acids across a collection of highly expressed mammalian protein-encoding genes, as for example disclosed by Seed (*supra*). Nor is it coterminous with a ranking of translational efficiency values obtained from an analysis of the translational efficiencies of codons in specific cell types, as disclosed for example in WO 99/02694 for COS-1 cells and epithelial cells and in WO 2004/024915 for CHO cells. Indeed, the present inventors have determined that codon modification of wild-type antigen-encoding polynucleotides to replace codons found in the wild-type sequence with codons having a higher preference for producing an immune response than the codons they replaced significantly enhances the immune response to the encoded antigen, as compared to the immune response obtained with the wild-type sequence. As a result, the present invention enables for the first time the construction of antigen-encoding polynucleotides, which are codon-optimized for efficient production of immune responses in a mammal.

[0010] Thus, in one aspect of the present invention, methods are provided for constructing a synthetic polynucleotide from which a polypeptide is producible to confer an immune response to a target antigen in a mammal in a different quality than that conferred by a parent polynucleotide that encodes the same polypeptide, wherein the polypeptide corresponds to at least a portion of the target antigen. These methods generally comprise: (a) selecting a first codon of the parent polynucleotide for replacement with a synonymous codon, wherein the synonymous codon is selected on the basis that it exhibits a different preference for conferring an immune response ("an immune response preference") than the first codon in a comparison of immune response preferences; and (b) replacing the first codon with the synonymous codon to construct the synthetic polynucleotide, wherein the comparison of immune response preferences of the codons is represented by TABLE 1:

TABLE 1

Amino Acid	Ranking of Immune Response Preferences for Synonymous Codons
Ala	$Ala^{GCT} > Ala^{GCC} > (Ala^{GCA}, Ala^{GCG})$
Arg	$(Arg^{CGA}, Arg^{CGC}, Arg^{CGT}, Arg^{AGA}) > (Arg^{AGG}, Arg^{CGG})$
Asn	$Asn^{AAC} > Asn^{AAT}$
Asp	$Asp^{GAC} > Asp^{GAT}$
Cys	$Cys^{TGC} > Cys^{TGT}$
Glu	$Glu^{GAA} > Glu^{GAG}$
Gln	$Gln^{CAA} = Gln^{CAG}$
Gly	$Gly^{GGA} > (Gly^{GGG}, Gly^{GGT}, Gly^{GGC})$
His	$His^{CAC} = His^{CAT}$
Ile	$Ile^{ATC} >> Ile^{ATT} > Ile^{ATA}$
Leu	$(Leu^{CTG}, Leu^{CTC}) > (Leu^{CTA}, Leu^{CTT}) >> Leu^{TTG} > Leu^{TTA}$
Lys	$Lys^{AAG} = Lys^{AAA}$
Phe	$Phe^{TTT} > Phe^{TTC}$
Pro	$Pro^{CCC} > Pro^{CCT} >> (Pro^{CCA}, Pro^{CCG})$
Ser	$Ser^{TCG} >> (Ser^{TCT}, Ser^{TCA}, Ser^{TCC}) >> (Ser^{AGC}, Ser^{AGT})$
Thr	$Thr^{ACG} > Thr^{ACC} >> Thr^{ACA} > Thr^{ACT}$
Tyr	$Tyr^{TAC} > Tyr^{TAT}$
Val	$(Val^{GTG}, Val^{GTC}) > Val^{GTT} > Val^{GTA}$

[0011] Thus, a stronger or enhanced immune response to the target antigen (*e.g.*, an immune response that is at least about 110%, 150%, 200%, 300%, 400%, 500%, 600%, 700%, 800%, 900%, 1000% and all integer percentages in between, of that produced from the parent polynucleotide under identical conditions) can be achieved by selecting a synonymous codon that has a higher immune response preference than the first codon it replaces. In

specific embodiments, the synonymous codon is selected such that it has a higher immune response preference that is at least about 10% (and at least about 11% to at least about 1000% and all integer percentages in between) higher than the immune response preference of the codon it replaces. In illustrative examples of this type, the first and synonymous codons are selected from TABLE 2:

TABLE 2

First Codon	Synonymous Codon	First Codon	Synonymous Codon	First Codon	Synonymous Codon
Ala^{GCG}	Ala^{GCT}	Ile^{ATA}	Ile^{ATC}	Ser^{AGT}	Ser^{TCG}
Ala^{GCG}	Ala^{GCC}	Ile^{ATA}	Ile^{ATT}	Ser^{AGT}	Ser^{TCT}
Ala^{GCA}	Ala^{GCT}	Ile^{ATT}	Ile^{ATC}	Ser^{AGT}	Ser^{TCA}
Ala^{GCA}	Ala^{GCC}			Ser^{AGT}	Ser^{TCC}
Ala^{GCC}	Ala^{GCT}	Leu^{TTA}	Leu^{CTG}	Ser^{AGC}	Ser^{TCG}
		Leu^{TTA}	Leu^{CTC}	Ser^{AGC}	Ser^{TCT}
Arg^{CGG}	Arg^{CGA}	Leu^{TTA}	Leu^{CTA}	Ser^{AGC}	Ser^{TCA}
Arg^{CGG}	Arg^{CGC}	Leu^{TTA}	Leu^{CTT}	Ser^{AGC}	Ser^{TCC}
Arg^{CGG}	Arg^{CGT}	Leu^{TTA}	Leu^{TTG}	Ser^{TCC}	Ser^{TCG}
Arg^{CGG}	Arg^{AGA}	Leu^{TTG}	Leu^{CTG}	Ser^{TCA}	Ser^{TCG}
Arg^{AGG}	Arg^{CGA}	Leu^{TTG}	Leu^{CTC}	Ser^{TCT}	Ser^{TCG}
Arg^{AGG}	Arg^{CGC}	Leu^{TTG}	Leu^{CTA}		
Arg^{AGG}	Arg^{CGT}	Leu^{TTG}	Leu^{CTT}	Thr^{ACT}	Thr^{ACG}
Arg^{AGG}	Arg^{AGA}	Leu^{CTT}	Leu^{CTG}	Thr^{ACT}	Thr^{ACC}
		Leu^{CTT}	Leu^{CTC}	Thr^{ACT}	Thr^{ACA}
Asn^{AAT}	Asn^{AAC}	Leu^{CTA}	Leu^{CTG}	Thr^{ACA}	Thr^{ACG}
		Leu^{CTA}	Leu^{CTC}	Thr^{ACA}	Thr^{ACC}
Asp^{GAT}	Asp^{GAC}			Thr^{ACC}	Thr^{ACG}
		Phe^{TTC}	Phe^{TTT}		

First Codon	Synonymous Codon	First Codon	Synonymous Codon	First Codon	Synonymous Codon
Cys^{TGT}	Cys^{TGC}			Tyr^{TAT}	Tyr^{TAC}
		Pro^{CCG}	Pro^{CCC}		
Glu^{GAG}	Glu^{GAA}	Pro^{CCG}	Pro^{CCT}	Val^{GTA}	Val^{GTG}
		Pro^{CCA}	Pro^{CCC}	Val^{GTA}	Val^{GTC}
Gly^{GGC}	Gly^{GGA}	Pro^{CCA}	Pro^{CCT}	Val^{GTA}	Val^{GTT}
Gly^{GGT}	Gly^{GGA}	Pro^{CCT}	Pro^{CCC}	Val^{GTT}	Val^{GTG}
Gly^{GGG}	Gly^{GGA}			Val^{GTT}	Val^{GTC}

[0012] In other illustrative examples of this type, the first and synonymous codons are selected from TABLE 3:

TABLE 3

First Codon	Synonymous Codon	First Codon	Synonymous Codon	First Codon	Synonymous Codon
Ala^{GCG}	Ala^{GCT}	Leu^{TTA}	Leu^{CTA}	Ser^{AGT}	Ser^{TCG}
Ala^{GCA}	Ala^{GCT}	Leu^{TTA}	Leu^{CTT}	Ser^{AGT}	Ser^{TCT}
Ala^{GCC}	Ala^{GCT}	Leu^{TTA}	Leu^{TTG}	Ser^{AGT}	Ser^{TCA}
		Leu^{TTG}	Leu^{CTA}	Ser^{AGC}	Ser^{TCG}
Arg^{CGG}	Arg^{CGA}	Leu^{TTG}	Leu^{CTT}	Ser^{AGC}	Ser^{TCT}
Arg^{CGG}	Arg^{CGT}			Ser^{AGC}	Ser^{TCA}
Arg^{CGG}	Arg^{AGA}	Phe^{TTC}	Phe^{TTT}	Ser^{AGC}	Ser^{TCC}
Arg^{AGG}	Arg^{CGA}			Ser^{TCC}	Ser^{TCG}
Arg^{AGG}	Arg^{CGT}	Pro^{CCG}	Pro^{CCT}	Ser^{TCA}	Ser^{TCG}
Arg^{AGG}	Arg^{AGA}	Pro^{CCA}	Pro^{CCT}	Ser^{TCT}	Ser^{TCG}

First Codon	Synonymous Codon	First Codon	Synonymous Codon	First Codon	Synonymous Codon
Glu^{GAG}	Glu^{GAA}			Thr^{ACT}	Thr^{ACG}
				Thr^{ACT}	Thr^{ACA}
Gly^{GGC}	Gly^{GGA}			Thr^{ACA}	Thr^{ACG}
Gly^{GGT}	Gly^{GGA}			Thr^{ACC}	Thr^{ACG}
Gly^{GGG}	Gly^{GGA}				
				Val^{GTA}	Val^{GTT}

[0013] Suitably, in some of the illustrative examples noted above, the method further comprises selecting a second codon of the parent polynucleotide for replacement with a synonymous codon, wherein the synonymous codon is selected on the basis that it exhibits a higher immune response preference than the second codon in a comparison of immune response preferences; and (b) replacing the second codon with the synonymous codon, wherein the comparison of immune response preferences of the codons is represented by TABLE 4:

TABLE 4

Second Codon	Synonymous Codon	Second Codon	Synonymous Codon	Second Codon	Synonymous Codon
Ala^{GCG}	Ala^{GCT}	Ile^{ATA}	Ile^{ATC}	Ser^{AGT}	Ser^{TCG}
Ala^{GCG}	Ala^{GCC}	Ile^{ATA}	Ile^{ATT}	Ser^{AGT}	Ser^{TCT}
Ala^{GCA}	Ala^{GCT}	Ile^{ATT}	Ile^{ATC}	Ser^{AGT}	Ser^{TCA}
Ala^{GCA}	Ala^{GCC}			Ser^{AGT}	Ser^{TCC}
Ala^{GCC}	Ala^{GCT}	Leu^{TTA}	Leu^{CTG}	Ser^{AGC}	Ser^{TCG}
		Leu^{TTA}	Leu^{CTC}	Ser^{AGC}	Ser^{TCT}
Arg^{CGG}	Arg^{CGA}	Leu^{TTA}	Leu^{CTA}	Ser^{AGC}	Ser^{TCA}
Arg^{CGG}	Arg^{CGC}	Leu^{TTA}	Leu^{CTT}	Ser^{AGC}	Ser^{TCC}

Second Codon	Synonymous Codon	Second Codon	Synonymous Codon	Second Codon	Synonymous Codon
Arg^{CGG}	Arg^{CGT}	Leu^{TTA}	Leu^{TTG}	Ser^{TCC}	Ser^{TCG}
Arg^{CGG}	Arg^{AGA}	Leu^{TTG}	Leu^{CTG}	Ser^{TCA}	Ser^{TCG}
Arg^{AGG}	Arg^{CGA}	Leu^{TTG}	Leu^{CTC}	Ser^{TCT}	Ser^{TCG}
Arg^{AGG}	Arg^{CGC}	Leu^{TTG}	Leu^{CTA}		
Arg^{AGG}	Arg^{CGT}	Leu^{TTG}	Leu^{CTT}	Thr^{ACT}	Thr^{ACG}
Arg^{AGG}	Arg^{AGA}	Leu^{CTT}	Leu^{CTG}	Thr^{ACT}	Thr^{ACC}
		Leu^{CTT}	Leu^{CTC}	Thr^{ACT}	Thr^{ACA}
Asn^{AAT}	Asn^{AAC}	Leu^{CTA}	Leu^{CTG}	Thr^{ACA}	Thr^{ACG}
		Leu^{CTA}	Leu^{CTC}	Thr^{ACA}	Thr^{ACC}
Asp^{GAT}	Asp^{GAC}			Thr^{ACC}	Thr^{ACG}
		Phe^{TTC}	Phe^{TTT}		
Cys^{TGT}	Cys^{TGC}			Tyr^{TAT}	Tyr^{TAC}
		Pro^{CCG}	Pro^{CCC}		
Glu^{GAG}	Glu^{GAA}	Pro^{CCG}	Pro^{CCT}	Val^{GTA}	Val^{GTG}
		Pro^{CCA}	Pro^{CCC}	Val^{GTA}	Val^{GTC}
Gly^{GGC}	Gly^{GGA}	Pro^{CCA}	Pro^{CCT}	Val^{GTA}	Val^{GTT}
Gly^{GGT}	Gly^{GGA}	Pro^{CCT}	Pro^{CCC}	Val^{GTT}	Val^{GTG}
Gly^{GGG}	Gly^{GGA}			Val^{GTT}	Val^{GTC}

[0014] Conversely, a weaker or reduced immune response to the target antigen (*e.g.*, an immune response that is at less than about 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%, 1% and all integer percentages in between, of that produced from the parent polynucleotide under identical conditions) can be achieved by selecting a synonymous codon that has a lower immune response preference than the first codon it replaces. In specific embodiments of this type, the synonymous codon is selected such that it has an immune

response preference that is less than about 90% of the immune response preference of the codon it replaces. In illustrative examples, the first and synonymous codons are selected from the TABLE 5:

TABLE 5

First Codon	Synonymous Codon	First Codon	Synonymous Codon	First Codon	Synonymous Codon
Ala^{GCT}	Ala^{GCG}	Ile^{ATC}	Ile^{ATA}	Ser^{TCG}	Ser^{TCT}
Ala^{GCT}	Ala^{GCA}	Ile^{ATC}	Ile^{ATT}	Ser^{TCG}	Ser^{TCA}
Ala^{GCT}	Ala^{GCC}	Ile^{ATT}	Ile^{ATA}	Ser^{TCG}	Ser^{TCC}
Ala^{GCC}	Ala^{GCG}			Ser^{TCG}	Ser^{AGC}
Ala^{GCC}	Ala^{GCA}	Leu^{CTG}	Leu^{CTA}	Ser^{TCG}	Ser^{AGT}
		Leu^{CTG}	Leu^{CTT}	Ser^{TCT}	Ser^{AGC}
Arg^{CGA}	Arg^{AGG}	Leu^{CTG}	Leu^{TTG}	Ser^{TCT}	Ser^{AGT}
Arg^{CGA}	Arg^{CGG}	Leu^{CTG}	Leu^{TTA}	Ser^{TCA}	Ser^{AGC}
Arg^{CGC}	Arg^{AGG}	Leu^{CTC}	Leu^{CTA}	Ser^{TCA}	Ser^{AGT}
Arg^{CGC}	Arg^{CGG}	Leu^{CTC}	Leu^{CTT}	Ser^{TCC}	Ser^{AGC}
Arg^{CGT}	Arg^{AGG}	Leu^{CTC}	Leu^{TTG}	Ser^{TCC}	Ser^{AGT}
Arg^{CGT}	Arg^{CGG}	Leu^{CTC}	Leu^{TTA}		
Arg^{AGA}	Arg^{AGG}	Leu^{CTA}	Leu^{TTG}	Thr^{ACG}	Thr^{ACC}
Arg^{AGA}	Arg^{CGG}	Leu^{CTA}	Leu^{TTA}	Thr^{ACG}	Thr^{ACA}
		Leu^{CTT}	Leu^{TTG}	Thr^{ACG}	Thr^{ACT}
Asn^{AAC}	Asn^{AAT}	Leu^{CTT}	Leu^{TTA}	Thr^{ACC}	Thr^{ACA}
		Leu^{TTG}	Leu^{TTA}	Thr^{ACC}	Thr^{ACT}
Asp^{GAC}	Asp^{GAT}			Thr^{ACA}	Thr^{ACT}
		Phe^{TTT}	Phe^{TTC}		
Cys^{TGC}	Cys^{TGT}			Tyr^{TAC}	Tyr^{TAT}
		Pro^{CCC}	Pro^{CCT}		

First Codon	Synonymous Codon	First Codon	Synonymous Codon	First Codon	Synonymous Codon
Glu^{GAA}	Glu^{GAG}	Pro^{CCC}	Pro^{CCA}	Val^{GTG}	Val^{GTT}
		Pro^{CCC}	Pro^{CCG}	Val^{GTG}	Val^{GTA}
Gly^{GGA}	Gly^{GGC}	Pro^{CCT}	Pro^{CCA}	Val^{GTC}	Val^{GTT}
Gly^{GGA}	Gly^{GGT}	Pro^{CCT}	Pro^{CCG}	Val^{GTC}	Val^{GTA}
Gly^{GGA}	Gly^{GGG}			Val^{GTT}	Val^{GTA}

[0015] In other illustrative examples, the first and synonymous codons are selected from TABLE 6:

TABLE 6

First Codon	Synonymous Codon	First Codon	Synonymous Codon	First Codon	Synonymous Codon
Ala^{GCT}	Ala^{GCG}	Leu^{CTA}	Leu^{TTG}	Ser^{TCG}	Ser^{TCT}
Ala^{GCT}	Ala^{GCA}	Leu^{CTA}	Leu^{TTA}	Ser^{TCG}	Ser^{TCA}
Ala^{GCT}	Ala^{GCC}	Leu^{CTT}	Leu^{TTG}	Ser^{TCG}	Ser^{TCC}
		Leu^{CTT}	Leu^{TTA}	Ser^{TCG}	Ser^{AGC}
Arg^{CGA}	Arg^{AGG}	Leu^{TTG}	Leu^{TTA}	Ser^{TCG}	Ser^{AGT}
Arg^{CGA}	Arg^{CGG}			Ser^{TCT}	Ser^{AGC}
Arg^{CGT}	Arg^{AGG}	Phe^{TTT}	Phe^{TTC}	Ser^{TCT}	Ser^{AGT}
Arg^{CGT}	Arg^{CGG}			Ser^{TCA}	Ser^{AGC}
Arg^{AGA}	Arg^{AGG}	Pro^{CCT}	Pro^{CCA}	Ser^{TCA}	Ser^{AGT}
Arg^{AGA}	Arg^{CGG}	Pro^{CCT}	Pro^{CCG}	Ser^{TCC}	Ser^{AGC}
Glu^{GAA}	Glu^{GAG}			Thr^{ACG}	Thr^{ACC}
				Thr^{ACG}	Thr^{ACA}

First Codon	Synonymous Codon	First Codon	Synonymous Codon	First Codon	Synonymous Codon
Gly^{GGA}	Gly^{GGC}			Thr^{ACG}	Thr^{ACT}
Gly^{GGA}	Gly^{GGT}			Thr^{ACA}	Thr^{ACT}
Gly^{GGA}	Gly^{GGG}				
				Val^{GTT}	Val^{GTA}

[0016] Suitably, in some of the illustrative examples noted above, the method further comprises selecting a second codon of the parent polynucleotide for replacement with a synonymous codon, wherein the synonymous codon is selected on the basis that it exhibits a lower immune response preference than the second codon in a comparison of immune response preferences; and; (b) replacing the second codon with the synonymous codon, wherein the comparison of immune response preferences of the codons is represented by TABLE 7:

TABLE 7

Second Codon	Synonymous Codon	Second Codon	Synonymous Codon	Second Codon	Synonymous Codon
Ala^{GCT}	Ala^{GCG}	Ile^{ATC}	Ile^{ATA}	Ser^{TCG}	Ser^{TCT}
Ala^{GCT}	Ala^{GCA}	Ile^{ATC}	Ile^{ATT}	Ser^{TCG}	Ser^{TCA}
Ala^{GCT}	Ala^{GCC}	Ile^{ATT}	Ile^{ATA}	Ser^{TCG}	Ser^{TCC}
Ala^{GCC}	Ala^{GCG}			Ser^{TCG}	Ser^{AGC}
Ala^{GCC}	Ala^{GCA}	Leu^{CTG}	Leu^{CTA}	Ser^{TCG}	Ser^{AGT}
		Leu^{CTG}	Leu^{CTT}	Ser^{TCT}	Ser^{AGC}
Arg^{CGA}	Arg^{AGG}	Leu^{CTG}	Leu^{TTG}	Ser^{TCT}	Ser^{AGT}
Arg^{CGA}	Arg^{CGG}	Leu^{CTG}	Leu^{TTA}	Ser^{TCA}	Ser^{AGC}
Arg^{CGC}	Arg^{AGG}	Leu^{CTC}	Leu^{CTA}	Ser^{TCA}	Ser^{AGT}
Arg^{CGC}	Arg^{CGG}	Leu^{CTC}	Leu^{CTT}	Ser^{TCC}	Ser^{AGC}

Second Codon	Synonymous Codon	Second Codon	Synonymous Codon	Second Codon	Synonymous Codon
Arg^{CGT}	Arg^{AGG}	Leu^{CTC}	Leu^{TTG}	Ser^{TCC}	Ser^{AGT}
Arg^{CGT}	Arg^{CGG}	Leu^{CTC}	Leu^{TTA}		
Arg^{AGA}	Arg^{AGG}	Leu^{CTA}	Leu^{TTG}	Thr^{ACG}	Thr^{ACC}
Arg^{AGA}	Arg^{CGG}	Leu^{CTA}	Leu^{TTA}	Thr^{ACG}	Thr^{ACA}
		Leu^{CTT}	Leu^{TTG}	Thr^{ACG}	Thr^{ACT}
Asn^{AAC}	Asn^{AAT}	Leu^{CTT}	Leu^{TTA}	Thr^{ACC}	Thr^{ACA}
		Leu^{TTG}	Leu^{TTA}	Thr^{ACC}	Thr^{ACT}
Asp^{GAC}	Asp^{GAT}			Thr^{ACA}	Thr^{ACT}
		Phe^{TTT}	Phe^{TTC}		
Cys^{TGC}	Cys^{TGT}			Tyr^{TAC}	Tyr^{TAT}
		Pro^{CCC}	Pro^{CCT}		
Glu^{GAA}	Glu^{GAG}	Pro^{CCC}	Pro^{CCA}	Val^{GTG}	Val^{GTT}
		Pro^{CCC}	Pro^{CCG}	Val^{GTG}	Val^{GTA}
Gly^{GGA}	Gly^{GGC}	Pro^{CCT}	Pro^{CCA}	Val^{GTC}	Val^{GTT}
Gly^{GGA}	Gly^{GGT}	Pro^{CCT}	Pro^{CCG}	Val^{GTC}	Val^{GTA}
Gly^{GGA}	Gly^{GGG}			Val^{GTT}	Val^{GTA}

[0017] In another aspect, the invention provides a synthetic polynucleotide constructed according to any one of the above methods.

[0018] In accordance with the present invention, synthetic polynucleotides that are constructed by methods described herein are useful for expression in a mammal to elicit an immune response to a target antigen. Accordingly, in yet another aspect, the present invention provides chimeric constructs that comprise a synthetic polynucleotide of the invention, which is operably connected to a regulatory polynucleotide.

[0019] In some embodiments, the chimeric construct is in the form of a pharmaceutical composition that optionally comprises a pharmaceutically acceptable excipient and/or carrier. Accordingly, in another aspect, the invention provides pharmaceutical compositions that are useful for modulating an immune response to a target antigen in a mammal, which response is conferred by the expression of a parent polynucleotide that encodes a polypeptide corresponding to at least a portion of the target antigen. These compositions generally comprise a chimeric construct and a pharmaceutically acceptable excipient and/or carrier, wherein the chimeric construct comprises a synthetic polynucleotide that is operably connected to a regulatory polynucleotide and that is distinguished from the parent polynucleotide by the replacement of a first codon in the parent polynucleotide with a synonymous codon that has a different immune response preference than the first codon and wherein the first and synonymous codons are selected according to any one of TABLES 2, 3, 5 and 6. In some embodiments, the compositions further comprise an adjuvant that enhances the effectiveness of the immune response. In some embodiments, the composition is formulated for transcutaneous or dermal administration, *e.g.*, by biolistic or microneedle delivery or by intradermal injection. Suitably, in embodiments in which a stronger or enhanced immune response to the target antigen is desired, the first and synonymous codons are selected according to TABLES 2 or 3. Conversely, in embodiments in which a weaker or reduced immune response to the target antigen is desired, the first and synonymous codons are selected according to TABLES 5 or 6.

[0020] In yet another aspect, the invention embraces methods of modulating the quality of an immune response to a target antigen in a mammal, which response is conferred by the expression of a parent polynucleotide that encodes a polypeptide corresponding to at least a portion of the target antigen. These methods generally comprise: introducing into the mammal a synthetic polynucleotide that is operably connected to a regulatory polynucleotide and that is distinguished from the parent polynucleotide by the replacement of a first codon in the parent polynucleotide with a synonymous codon that has a different immune response preference than the first codon and wherein the first and synonymous codons are selected according to any one of TABLES 2, 3, 5 and 6. In these methods, expression of the synthetic polynucleotide results in a different quality (*e.g.*, stronger or weaker) of immune response than the one obtained through expression of the parent polynucleotide under the same conditions. Suitably, the chimeric construct is introduced into the mammal by delivering the construct to antigen-presenting cells (*e.g.*, dendritic cells, macrophages, Langerhans cells or their precursors) of the mammal. In some embodiments, the chimeric construct is introduced

into the dermis and/or epidermis of the mammal (*e.g.*, by transcutaneous or intradermal administration) and in this regard any suitable administration site is envisaged including the abdomen. Generally, the immune response is selected from a cell-mediated response and a humoral immune response. In some embodiments, the immune response is a humoral immune response. In other embodiments, the immune response is a cellular immune response. In still other embodiments, the immune response is a humoral immune response and a cellular immune response.

[0021] In a related aspect, the invention encompasses methods of enhancing the quality of an immune response to a target antigen in a mammal, which response is conferred by the expression of a parent polynucleotide that encodes a polypeptide corresponding to at least a portion of the target antigen. These methods generally comprise: introducing into the mammal a chimeric construct comprising a synthetic polynucleotide that is operably connected to a regulatory polynucleotide and that is distinguished from the parent polynucleotide by the replacement of a first codon in the parent polynucleotide with a synonymous codon that has a higher immune response preference than the first codon, wherein the first and synonymous codons are selected according to TABLES 2 or 3. In these methods, expression of the synthetic polynucleotide typically results in a stronger or enhanced immune response than the one obtained through expression of the parent polynucleotide under the same conditions.

[0022] In another related aspect, the invention extends to methods of reducing the quality of an immune response to a target antigen in a mammal, which response is conferred by the expression of a parent polynucleotide that encodes a polypeptide corresponding to at least a portion of the target antigen. These methods generally comprise: introducing into the mammal a chimeric construct comprising a synthetic polynucleotide that is operably connected to a regulatory polynucleotide and that is distinguished from the parent polynucleotide by the replacement of a first codon in the parent polynucleotide with a synonymous codon that has a lower immune response preference than the first codon, wherein the first and synonymous codons are selected according to TABLES 5 or 6. In these methods, expression of the synthetic polynucleotide typically results in a weaker or reduced immune response than the one obtained through expression of the parent polynucleotide under the same conditions.

[0023] Yet a further aspect of the present invention embraces methods of enhancing the quality of an immune response to a target antigen in a mammal, which response is

conferred by the expression of a first polynucleotide that encodes a polypeptide corresponding to at least a portion of the target antigen. These methods generally comprise: co-introducing into the mammal a first nucleic acid construct comprising the first polynucleotide in operable connection with a regulatory polynucleotide; and a second nucleic acid construct comprising a second polynucleotide that is operably connected to a regulatory polynucleotide and that encodes an iso-tRNA corresponding to a codon of the first polynucleotide, wherein the codon has a low or intermediate immune response preference and is selected from the group consisting of Ala^{GCA}, Ala^{GCG}, Ala^{GCC}, Arg^{AGG}, Arg^{CGG}, Asn^{AAT}, Asp^{GAT}, Cys^{TGT}, Glu^{GAG}, Gly^{GGG}, Gly^{GGT}, Gly^{GGC}, Ile^{ATA}, Ile^{ATT}, Leu^{TTG}, Leu^{TTA}, Leu^{CTA}, Leu^{CTT}, Phe^{TTC}, Pro^{CCA}, Pro^{CCG}, Pro^{CCT}, Ser^{AGC}, Ser^{AGT}, Ser^{TCT}, Ser^{TCA}, Ser^{TCC}, Thr^{ACA}, Thr^{ACT}, Tyr^{TAT}, Val^{GTA} and Val^{GTT}. In specific embodiments, the codon has a 'low' immune response preference, and is selected from the group consisting of Ala^{GCA}, Ala^{GCG}, Arg^{AGG}, Arg^{CGG}, Asn^{AAT}, Asp^{GAT}, Cys^{TGT}, Glu^{GAG}, Gly^{GGG}, Gly^{GGT}, Gly^{GGC}, Ile^{ATA}, Leu^{TTG}, Leu^{TTA}, Phe^{TTC}, Pro^{CCA}, Pro^{CCG}, Ser^{AGC}, Ser^{AGT}, Thr^{ACT}, Tyr^{TAT} and Val^{GTA}.

BRIEF DESCRIPTION OF THE DRAWINGS

[0024] Figure 1 is a diagrammatic representation depicting a nucleotide sequence alignment of secreted ALA E7 constructs and controls (IgkC1, IgkS1-1, IgkS1-2, IgkS1-3, IgkS1-4 and IgkC2) as further defined in Example 1 and Table 12. The sequences are ligated into the *Kpn*I and *Eco*RI sites of pCDNA3.

[0025] Figure 2 is a diagrammatic representation depicting a nucleotide sequence alignment of secreted ARG E7 constructs and controls (IgkS1-5, IgkS1-6, IgkS1-7, IgkS1-8, IgkS1-9, IgkS1-10, IgkC1 and IgkC2) as further defined in Example 1 and Table 12. The sequences are ligated into the *Kpn*I and *Eco*RI sites of pCDNA3.

[0026] Figure 3 is a diagrammatic representation depicting a nucleotide sequence alignment of secreted ASN and LYS E7 constructs and controls (IgkC1, IgkS1-12, IgkS1-31 and IgkC2) as further defined in Example 1 and Table 12. The sequences are ligated into the *Kpn*I and *Eco*RI sites of pCDNA3.

[0027] Figure 4 is a diagrammatic representation depicting a nucleotide sequence alignment of secreted ASP E7 constructs and controls (IgkC1, IgkS1-13, IgkS1-14 and IgkC2) as further defined in Example 1 and Table 12. The sequences are ligated into the *Kpn*I and *Eco*RI sites of pCDNA3.

[0028] Figure 5 is a diagrammatic representation depicting a nucleotide sequence alignment of secreted CYS E7 constructs and controls (IgkC1, IgkS1-15, IgkS1-16 and IgkC2) as further defined in Example 1 and Table 12. The sequences are ligated into the *Kpn*I and *Eco*RI sites of pCDNA3.

[0029] Figure 6 is a diagrammatic representation depicting a nucleotide sequence alignment of secreted GLU E7 constructs and controls (IgkS1-17, IgkS1-18, IgkC2 and IgkC1) as further defined in Example 1 and Table 12. The sequences are ligated into the *Kpn*I and *Eco*RI sites of pCDNA3.

[0030] Figure 7 is a diagrammatic representation depicting a nucleotide sequence alignment of secreted GLN E7 constructs and controls (IgkC1, IgkS1-19, IgkS1-20 and IgkC2) as further defined in Example 1 and Table 12. The sequences are ligated into the *Kpn*I and *Eco*RI sites of pCDNA3.

[0031] Figure 8 is a diagrammatic representation depicting a nucleotide sequence alignment of secreted GLY E7 constructs and controls (IgkC1, IgkS1-21, IgkS1-22, IgkS1-23,

IgkS1-24 and IgkC2) as further defined in Example 1 and Table 12. The sequences are ligated into the *Kpn*I and *Eco*RI sites of pCDNA3.

[0032] Figure 9 is a diagrammatic representation depicting a nucleotide sequence alignment of secreted HIS E7 constructs and controls (IgkC1, IgkS1-25, IgkS1-26 and IgkC2) as further defined in Example 1 and Table 12. The sequences are ligated into the *Kpn*I and *Eco*RI sites of pCDNA3.

[0033] Figure 10 is a diagrammatic representation depicting a nucleotide sequence alignment of secreted ILE E7 constructs and controls (IgkC1, IgkS1-27, IgkS1-28, IgkS1-29 and IgkC2) as further defined in Example 1 and Table 12. The sequences are ligated into the *Kpn*I and *Eco*RI sites of pCDNA3.

[0034] Figure 11 is a diagrammatic representation depicting a nucleotide sequence alignment of secreted LEU E7 constructs and controls (IgkS1-50, IgkS1-51, IgkS1-52, IgkS1-53, IgkS1-54, IgkS1-55, IgkC3 and IgkC4) as further defined in Example 1 and Table 12. The sequences are ligated into the *Kpn*I and *Eco*RI sites of pCDNA3. The LEU E7 constructs are oncogenic (*i.e.*, encode wild-type E7 protein).

[0035] Figure 12 is a diagrammatic representation depicting a nucleotide sequence alignment of secreted PHE E7 constructs and controls (IgkS1-32, IgkS1-33, IgkC1and IgkC2) as further defined in Example 1 and Table 12. The sequences are ligated into the *Kpn*I and *Eco*RI sites of pCDNA3. Two LEU residues were mutated to PHE in this sequence so that there are three instead of one PHE residue.

[0036] Figure 13 is a diagrammatic representation depicting a nucleotide sequence alignment of secreted PRO E7 constructs and controls (IgkS1-56, IgkS1-57, IgkS1-58, IgkS1-59, IgkC3 and IgkC4) as further defined in Example 1 and Table 12. The sequences are ligated into the *Kpn*I and *Eco*RI sites of pCDNA3. The PRO E7 constructs are oncogenic (*i.e.*, encode wild-type E7 protein).

[0037] Figure 14 is a diagrammatic representation depicting a nucleotide sequence alignment of secreted SER E7 constructs and controls (IgkS1-34, IgkS1-35, IgkS1-36, IgkS1-37, IgkS1-38, IgkS1-39, IgkC1and IgkC2) as further defined in Example 1 and Table 12. The sequences are ligated into the *Kpn*I and *Eco*RI sites of pCDNA3.

[0038] Figure 15 is a diagrammatic representation depicting a nucleotide sequence alignment of secreted THR E7 constructs and controls (IgkC1, IgkS1-40, IgkS1-41, IgkS1-42,

IgkS1-43 and IgkC2) as further defined in Example 1 and Table 12. The sequences are ligated into the *Kpn*I and *Eco*RI sites of pCDNA3.

[0039] Figure 16 is a diagrammatic representation depicting a nucleotide sequence alignment of secreted TYR E7 constructs and controls (IgkC1, IgkS1-44, IgkS1-45 and IgkC2) as further defined in Example 1 and Table 12. The sequences are ligated into the *Kpn*I and *Eco*RI sites of pCDNA3.

[0040] Figure 17 is a diagrammatic representation depicting a nucleotide sequence alignment of secreted VAL E7 constructs and controls (IgkC1, IgkS1-46, IgkS1-47, IgkS1-48, IgkS1-49 and IgkC2) as further defined in Example 1 and Table 12. The sequences are ligated into the *Kpn*I and *Eco*RI sites of pCDNA3.

[0041] Figure 18 is a graphical representation showing the response to gene gun immunization with optimized and de-optimized E7 constructs measured by (a) ELISA, (b) Memory B cell ELISPOT, and (c) IFN-γ ELISPOT. For part (a) eight mice were immunized per group (4 immunizations, 3 weeks apart) and the sera taken three weeks after the final immunization; (left) E7 protein ELISA, (right) E7 peptide 101 ELISA. Wells were done in duplicate. For parts (b) and (c) mice were immunized twice, three weeks apart and the spleens collected three weeks after the second immunization. The spleens were pooled prior to analysis. The Memory B cell and IFN-γ ELISPOTs were conducted twice and three times, respectively, and the wells done in triplicate. Three mice were used per group per repeat. The results shown in parts (b) and (c) are from individual experiments and are representative of the complete data sets. The particular ELISPOT experimental data included here were gathered together with the corresponding data in Figure 20 and therefore may be directly compared. Unpaired two-tailed t-tests were used to compare the modified constructs to wild-type. *** $P<0.001$, ** $0.001 \leq P < 0.01$, * $0.01 \leq P \leq 0.05$, ns= not significant ($P>0.05$). In (a) O1-O3 were not significantly different from MC as measured by unpaired two-tailed t-tests. wt= wild-type codon usage E7; O1-O3= codon-optimized E7 constructs 1 to 3; W=codon de-optimized E7; MC= mammalian consensus codon usage E7.

[0042] Figure 19 is a graphical representation showing the response to immunization by intradermal injection with optimized and de-optimized constructs measured by (a) ELISA, (b) Memory B cell ELISPOT, and (c) IFN-γ ELISPOT. For part (a) eight mice were immunized per group (4 immunizations, 3 weeks apart) and the sera taken three weeks after the final immunization; (left) E7 protein ELISA, (right) E7 peptide 101 ELISA. Wells

were done in duplicate. For parts (b) and (c) mice were immunized twice, three weeks apart and the spleens collected three weeks after the second immunization. The spleens were pooled prior to analysis. The Memory B cell and IFN-γ ELISPOTs were conducted twice and three times, respectively, and the wells done in triplicate. Three mice were used per group per repeat. The results shown in parts (b) and (c) are from individual experiments and are representative of the complete data sets. The particular ELISPOT experimental data included here were gathered together with the corresponding data in Figure 20 and therefore may be directly compared. Unpaired two-tailed t-tests were used to compare the modified constructs to wild-type. *** $P<0.001$, ** $0.001 \leq P < 0.01$, * $0.01 \leq P \leq 0.05$, ns= not significant ($P>0.05$). In (a) O1-O3 were not significantly different from MC as measured by unpaired two-tailed t-tests. wt= wild-type codon usage E7; O1-O3= codon-optimized E7 constructs 1 to 3; W=codon de-optimized E7; MC= mammalian consensus codon usage E7.

[0043] Figure 20 is a graphical representation showing the results of an ELISA that measures binding of serum from mice immunized with various gD2 constructs by intradermal injection (white bars) or gene gun immunization (black bars), to C-terminally His-tagged gD2tr. Note that the His-tagged gD2tr protein was used in an unpurified state (in CHO cell supernatant) and that background readings of non-specific binding to control supernatant have been subtracted from the results.

TABLE 8

BRIEF DESCRIPTION OF THE SEQUENCES

SEQUENCE ID NUMBER	SEQUENCE	LENGTH
SEQ ID NO: 1	IgkS2-13 Asp GAT construct nucleotide sequence	387 nts
SEQ ID NO: 2	IgkS2-14 Asp GAC construct nucleotide sequence	387 nts
SEQ ID NO: 3	IgkS2-15 Cys TGT construct nucleotide sequence	387 nts
SEQ ID NO: 4	IgkS2-16 Cys TGC construct nucleotide sequence	387 nts
SEQ ID NO: 5	IgkS2-17 Glu GAG construct nucleotide sequence	387 nts
SEQ ID NO: 6	IgkS2- 18 Glu GAA construct nucleotide sequence	387 nts
SEQ ID NO: 7	IgkS2-19 Gln CAG construct nucleotide sequence	387 nts
SEQ ID NO: 8	IgkS2-20 Gln CAA construct nucleotide sequence	387 nts
SEQ ID NO: 9	IgkS2-21 Gly GGG construct nucleotide sequence	387 nts
SEQ ID NO: 10	IgkS2-22 Gly GGA construct nucleotide sequence	387 nts
SEQ ID NO: 11	IgkS2-23 Gly GGT construct nucleotide sequence	387 nts
SEQ ID NO: 12	IgkS2-24 Gly GGC construct nucleotide sequence	387 nts
SEQ ID NO: 13	IgkS2-27 Ile ATA construct nucleotide sequence	387 nts
SEQ ID NO: 14	IgkS2-28 Ile ATT construct nucleotide sequence	387 nts
SEQ ID NO: 15	IgkS2-29 Ile ATC construct nucleotide sequence	387 nts
SEQ ID NO: 16	IgkS2-34 Ser AGT construct nucleotide sequence	387 nts
SEQ ID NO: 17	IgkS2-35 Ser AGC construct nucleotide sequence	387 nts
SEQ ID NO: 18	IgkS2-36 Ser TCG construct nucleotide sequence	387 nts
SEQ ID NO: 19	IgkS2-37 Ser TCA construct nucleotide sequence	387 nts
SEQ ID NO: 20	IgkS2-38 Ser TCT construct nucleotide sequence	387 nts
SEQ ID NO: 21	IgkS2-39 Ser TCC construct nucleotide sequence	387 nts
SEQ ID NO: 22	IgkS2-40 Thr ACG construct nucleotide sequence	387 nts
SEQ ID NO: 23	IgkS2-41 Thr ACA construct nucleotide sequence	387 nts
SEQ ID NO: 24	IgkS2-42 Thr ACT construct nucleotide sequence	387 nts
SEQ ID NO: 25	IgkS2-43 Thr ACC construct nucleotide sequence	387 nts

SEQUENCE ID NUMBER	SEQUENCE	LENGTH
SEQ ID NO: 26	IgkS2-46 Val GTG construct nucleotide sequence	387 nts
SEQ ID NO: 27	IgkS2-47 Val GTA construct nucleotide sequence	387 nts
SEQ ID NO: 28	IgkS2-48 Val GTT construct nucleotide sequence	387 nts
SEQ ID NO: 29	IgkS2-49 Val GTG construct nucleotide sequence	387 nts
SEQ ID NO: 30	IgkS2-1 Ala GCG Linker nucleotide sequence	408 nts
SEQ ID NO: 31	IgkS2-2 Ala GCA Linker nucleotide sequence	408 nts
SEQ ID NO: 32	IgkS2-3 Ala GCT Linker nucleotide sequence	408 nts
SEQ ID NO: 33	IgkS2-4 Ala GCC Linker nucleotide sequence	408 nts
SEQ ID NO: 34	IgkS2-5 Arg AGG Linker nucleotide sequence	408 nts
SEQ ID NO: 35	IgkS2-6 Arg AGA Linker nucleotide sequence	408 nts
SEQ ID NO: 36	IgkS2-7 Arg CGG Linker nucleotide sequence	408 nts
SEQ ID NO: 37	IgkS2-8 Arg CGA Linker nucleotide sequence	408 nts
SEQ ID NO: 38	IgkS2-9 Arg CGT Linker nucleotide sequence	408 nts
SEQ ID NO: 39	IgkS2-10 Arg CGC Linker nucleotide sequence	408 nts
SEQ ID NO: 40	IgkS2-11 Asn AAT Linker nucleotide sequence	408 nts
SEQ ID NO: 41	IgkS2-12 Asn AAC Linker nucleotide sequence	408 nts
SEQ ID NO: 42	IgkS2-25 His CAT Linker nucleotide sequence	408 nts
SEQ ID NO: 43	IgkS2-26 His CAC Linker nucleotide sequence	408 nts
SEQ ID NO: 44	IgkS2-30 Lys AAG Linker nucleotide sequence	408 nts
SEQ ID NO: 45	IgkS2-31 Lys AAA Linker nucleotide sequence	408 nts
SEQ ID NO: 46	IgkS2-32 Phe TTT Linker nucleotide sequence	408 nts
SEQ ID NO: 47	IgkS2-33 Phe TTC Linker nucleotide sequence	408 nts
SEQ ID NO: 48	IgkS2-44 Tyr TAT Linker nucleotide sequence	408 nts
SEQ ID NO: 49	IgkS2-45 Tyr TAC Linker nucleotide sequence	408 nts
SEQ ID NO: 50	Influenza A Virus HA hemagglutinin (A/Hong Kong/213/03(H5N1)) BAE07201 **wild-type**	1707 nts
SEQ ID NO: 51	Influenza A Virus HA hemagglutinin (A/Hong Kong/213/03(H5N1)) BAE07201 **wild-type**	568 aa
SEQ ID NO: 52	Influenza A Virus HA hemagglutinin (A/Hong	1707 nts

SEQUENCE ID NUMBER	SEQUENCE	LENGTH
	Kong/213/03(H5N1)) **Codon modified**	
SEQ ID NO: 53	Influenza A Virus HA hemagglutinin (A/swine/Korea/PZ72-1/2006 (H3N1)) DQ923506 **wild-type**	1701 nts
SEQ ID NO: 54	Influenza A Virus HA hemagglutinin (A/swine/Korea/PZ72-1/2006 (H3N1)) DQ923506 **wild-type**	566 aa
SEQ ID NO: 55	Influenza A Virus HA hemagglutinin (A/swine/Korea/PZ72-1/2006 (H3N1)) **Codon modified**	1701 nts
SEQ ID NO: 56	Influenza A Virus NA neuraminidase (A/Hong Kong/213/03(H5N1)) AB212056 **wild-type**	1410 nts
SEQ ID NO: 57	Influenza A Virus NA neuraminidase (A/Hong Kong/213/03(H5N1)) AB212056 **wild-type**	469 aa
SEQ ID NO: 58	Influenza A Virus NA neuraminidase (A/Hong Kong/213/03(H5N1)) **Codon modified**	1410 nts
SEQ ID NO: 59	Influenza A Virus NA neuraminidase (A/swine/MI/PU243/04 (H3N1)) DQ150427 **wild-type**	1410 nts
SEQ ID NO: 60	Influenza A Virus NA neuraminidase (A/swine/MI/PU243/04 (H3N1)) DQ150427 **wild-type**	469 aa
SEQ ID NO: 61	Influenza A Virus NA neuraminidase (A/swine/MI/PU243/04 (H3N1)) **Codon modified**	1410 nts
SEQ ID NO: 62	Hepatitis C Virus E1 (Serotype 1A, isolate H77) AF009606 **wild-type**	576 nts
SEQ ID NO: 63	Hepatitis C Virus E1 (Serotype 1A, isolate H77) NP 751920 **wild-type**	192 aa
SEQ ID NO: 64	Hepatitis C Virus E1 (Serotype 1A, isolate H77) **Codon modified**	576 nts
SEQ ID NO: 65	Hepatitis C Virus E2 (Serotype 1A, isolate H77) AF009606 **wild-type**	1089 nts
SEQ ID NO: 66	Hepatitis C Virus E2 (Serotype 1A, isolate H77) NP 751921 **wild-type**	363 aa
SEQ ID NO: 67	Hepatitis C Virus E2 (Serotype 1A, isolate H77) **Codon modified**	1089 nts
SEQ ID NO: 68	Epstein Barr Virus (Type 1, gp350 B95-8) NC 007605 **wild-type**	2724 nts

SEQUENCE ID NUMBER	SEQUENCE	LENGTH
SEQ ID NO: 69	Epstein Barr Virus (Type 1, gp350 B95-8) CAD53417 **wild-type**	907 aa
SEQ ID NO: 70	Epstein Barr Virus (Type 1, gp350 B95-8) **Codon modified**	2724 nts
SEQ ID NO: 71	Epstein Barr Virus (Type 2, gp350 AG876) NC 009334 **wild-type**	2661 nts
SEQ ID NO: 72	Epstein Barr Virus (Type 2, gp350 AG876) YP 001129462 **wild-type**	886 aa
SEQ ID NO: 73	Epstein Barr Virus (Type 2, gp350 AG876) **Codon Modified**	2661 nts
SEQ ID NO: 74	Herpes Simplex Virus 2 (Glycoprotein B strain HG52) NC 001798 **wild-type**	2715 nts
SEQ ID NO: 75	Herpes Simplex Virus 2 (Glycoprotein B strain HG52) CAB06752 **wild-type**	904 aa
SEQ ID NO: 76	Herpes Simplex Virus 2 (Glycoprotein B strain HG52) **Codon modified**	2715 nts
SEQ ID NO: 77	Herpes Simplex Virus (Glycoprotein D strain HG52) NC 001798 **wild-type**	1182 nts
SEQ ID NO: 78	Herpes Simplex Virus (Glycoprotein D strain HG52) NP 0044536 **wild-type**	393 aa
SEQ ID NO: 79	Herpes Simplex Virus (Glycoprotein D strain HG52) **Codon modified**	1182 nts
SEQ ID NO: 80	HPV-16 E7 **wild-type**	387 nts
SEQ ID NO: 81	HPV-16 E7 **O1**	387 nts
SEQ ID NO: 82	HPV-16 E7 **O2**	387 nts
SEQ ID NO: 83	HPV-16 E7 **O3**	417 nts
SEQ ID NO: 84	HPV-16 E7 **W**	387 nts
SEQ ID NO: 85	HSV-2 gD2 **wild-type**	1182 nts
SEQ ID NO: 86	HSV-2 gD2 **O1**	1182 nts
SEQ ID NO: 87	HSV-2 gD2 **O2**	1182 nts
SEQ ID NO: 88	HSV-2 gD2 **O3**	1182 nts
SEQ ID NO: 89	HSV-2 gD2 **W**	1182 nts

SEQUENCE ID NUMBER	SEQUENCE	LENGTH
SEQ ID NO: 90	Common forward primer	41 nts
SEQ ID NO: 91	ODN-7909	24 nts

DETAILED DESCRIPTION OF THE INVENTION

1. Definitions

[0044] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by those of ordinary skill in the art to which the invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, preferred methods and materials are described. For the purposes of the present invention, the following terms are defined below.

[0045] The articles "a" and "an" are used herein to refer to one or to more than one (*i.e.* to at least one) of the grammatical object of the article. By way of example, "an element" means one element or more than one element.

[0046] By "about" is meant a quantity, level, value, frequency, percentage, dimension, size, or amount that varies by no more than 15%, and preferably by no more than 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1% to a reference quantity, level, value, frequency, percentage, dimension, size, or amount.

[0047] The terms "administration concurrently" or "administering concurrently" or "co-administering" and the like refer to the administration of a single composition containing two or more actives, or the administration of each active as separate compositions and/or delivered by separate routes either contemporaneously or simultaneously or sequentially within a short enough period of time that the effective result is equivalent to that obtained when all such actives are administered as a single composition. By "simultaneously" is meant that the active agents are administered at substantially the same time, and desirably together in the same formulation. By "contemporaneously" it is meant that the active agents are administered closely in time, *e.g.*, one agent is administered within from about one minute to within about one day before or after another. Any contemporaneous time is useful. However, it will often be the case that when not administered simultaneously, the agents will be administered within about one minute to within about eight hours and preferably within less than about one to about four hours. When administered contemporaneously, the agents are suitably administered at the same site on the subject. The term "same site" includes the exact location, but can be within about 0.5 to about 15 centimeters, preferably from within about 0.5 to about 5 centimeters. The term "separately" as used herein means that the agents are administered at an interval, for example at an interval of about a day to several weeks or

months. The active agents may be administered in either order. The term "sequentially" as used herein means that the agents are administered in sequence, for example at an interval or intervals of minutes, hours, days or weeks. If appropriate the active agents may be administered in a regular repeating cycle.

[0048] As used herein, the term "*cis*-acting sequence" or "*cis*-regulatory region" or similar term shall be taken to mean any sequence of nucleotides which is derived from an expressible genetic sequence wherein the expression of the genetic sequence is regulated, at least in part, by the sequence of nucleotides. Those skilled in the art will be aware that a cis-regulatory region may be capable of activating, silencing, enhancing, repressing or otherwise altering the level of expression and/or cell-type-specificity and/or developmental specificity of any structural gene sequence.

[0049] Throughout this specification, unless the context requires otherwise, the words "comprise," "comprises" and "comprising" will be understood to imply the inclusion of a stated step or element or group of steps or elements but not the exclusion of any other step or element or group of steps or elements.

[0050] As used herein, a "chimeric construct" refers to a polynucleotide having heterologous nucleic acid elements. Chimeric constructs include "expression cassettes" or "expression constructs," which refer to an assembly that is capable of directing the expression of the sequence(s) or gene(s) of interest. An expression cassette generally includes control elements such as a promoter that is operably linked to (so as to direct transcription of) a synthetic polynucleotide of the invention, and often includes a polyadenylation sequence as well. Within certain embodiments of the invention, the chimeric construct may be contained within a vector. In addition to the components of the chimeric construct, the vector may include, one or more selectable markers, a signal which allows the vector to exist as single-stranded DNA (*e.g.*, a M13 origin of replication), at least one multiple cloning site, and a "mammalian" origin of replication (*e.g.*, a SV40 or adenovirus origin of replication).

[0051] As used herein, "conferred immune response," "immune response that is conferred" and the like refer to a temporary or permanent change in immune response to a target antigen, which occurs or would occur after the introduction of a polynucleotide to the mammal, and which would not occur in the absence of that introduction. Typically, such a temporary or permanent change occurs as a result of the transcription and/or translation of genetic information contained within that polynucleotide in a cell, or in at least one cell or cell type or class of cell within a mammal or within a class of mammals, and can be used to

distinguish the mammal, or class of mammals to which the polynucleotide has been provided from a similar mammal, or class of mammals, to which the polynucleotide has not been provided.

[0052] By "corresponds to" or "corresponding to" is meant an antigen which encodes an amino acid sequence that displays substantial similarity to an amino acid sequence in a target antigen. In general the antigen will display at least about 30, 40, 50, 55, 60, 65, 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 % similarity or identity to at least a portion of the target antigen (*e.g.*, at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or 95% of the amino acid sequence of the target antigen).

[0053] By "effective amount," in the context of modulating an immune response or treating or preventing a disease or condition, is meant the administration of that amount of composition to an individual in need thereof, either in a single dose or as part of a series, that is effective for achieving that modulation, treatment or prevention. The effective amount will vary depending upon the health and physical condition of the individual to be treated, the taxonomic group of individual to be treated, the formulation of the composition, the assessment of the medical situation, and other relevant factors. It is expected that the amount will fall in a relatively broad range that can be determined through routine trials.

[0054] The terms "enhancing an immune response," "producing a stronger immune response" and the like refer to increasing an animal's capacity to respond to a target antigen (*e.g.*, a foreign or disease-specific antigen or a self antigen), which can be determined for example by detecting an increase in the number, activity, and ability of the animal's cells that are primed to attack such antigens or an increase in the titer or activity of antibodies in the animal, which are immuno-interactive with the target antigen. Strength of immune response can be measured by standard immunoassays including: direct measurement of antibody titers or peripheral blood lymphocytes; cytolytic T lymphocyte assays; assays of natural killer cell cytotoxicity; cell proliferation assays including lymphoproliferation (lymphocyte activation) assays; immunoassays of immune cell subsets; assays of T-lymphocytes specific for the antigen in a sensitized subject; skin tests for cell-mediated immunity; etc. Such assays are well known in the art. See, *e.g.*, Erickson *et al.*, 1993, J. Immunol. 151:4189-4199; Doe *et al.*, 1994, Eur. J. Immunol. 24:2369-2376. Recent methods of measuring cell-mediated immune response include measurement of intracellular cytokines or cytokine secretion by T-cell populations, or by measurement of epitope specific T-cells (*e.g.*, by the tetramer technique) (reviewed by McMichael, A. J., and O'Callaghan, C. A., 1998, J. Exp. Med. 187(9)1367-

1371; Mcheyzer-Williams, M. G., *et al.*, 1996, Immunol. Rev. 150:5-21; Lalvani, A., *et al.*, 1997, J. Exp. Med. 186:859-865). Any statistically significant increase in strength of immune response as measured for example by immunoassay is considered an "enhanced immune response" or "immunoenhancement" as used herein. Enhanced immune response is also indicated by physical manifestations such as fever and inflammation, as well as healing of systemic and local infections, and reduction of symptoms in disease, *i.e.*, decrease in tumor size, alleviation of symptoms of a disease or condition including, but not restricted to, leprosy, tuberculosis, malaria, naphthous ulcers, herpetic and papillomatous warts, gingivitis, arthrosclerosis, the concomitants of AIDS such as Kaposi's sarcoma, bronchial infections, and the like. Such physical manifestations also encompass "enhanced immune response" or "immunoenhancement" as used herein. By contrast, "reducing an immune response," "producing a weaker immune response" and the like refer to decreasing an animal's capacity to respond to a target antigen, which can be determined for example by conducting immunoassays or assessing physical manifestations, as described for example above.

[0055] The terms "expression" or "gene expression" refer to production of RNA message and/or translation of RNA message into proteins or polypeptides.

[0056] By "expression vector" is meant any autonomous genetic element capable of directing the synthesis of a protein encoded by the vector. Such expression vectors are known by practitioners in the art.

[0057] The term "gene" is used in its broadest context to include both a genomic DNA region corresponding to the gene as well as a cDNA sequence corresponding to exons or a recombinant molecule engineered to encode a functional form of a product.

[0058] As used herein the term "heterologous" refers to a combination of elements that are not naturally occurring or that are obtained from different sources.

[0059] "Immune response" or "immunological response" refers to the concerted action of lymphocytes, antigen-presenting cells, phagocytic cells, granulocytes, and soluble macromolecules produced by the above cells or the liver (including antibodies, cytokines, and complement) that results in selective damage to, destruction of, or elimination from the body of cancerous cells, metastatic tumor cells, metastatic breast cancer cells, invading pathogens, cells or tissues infected with pathogens, or, in cases of autoimmunity or pathological inflammation, normal human cells or tissues. In some embodiments, an "immune response' encompasses the development in an individual of a humoral and/or a cellular immune

response to a polypeptide that is encoded by an introduced synthetic polynucleotide of the invention. As known in the art, the terms "humoral immune response" includes and encompasses an immune response mediated by antibody molecules, while a "cellular immune response" includes and encompasses an immune response mediated by T-lymphocytes and/or other white blood cells. Thus, an immune response that is stimulated by a synthetic polynucleotide of the invention may be one that stimulates the production of antibodies (*e.g.*, neutralizing antibodies that block bacterial toxins and pathogens such as viruses entering cells and replicating by binding to toxins and pathogens, typically protecting cells from infection and destruction). The synthetic polynucleotide may also elicit production of cytolytic T lymphocytes (CTLs). Hence, an immunological response may include one or more of the following effects: the production of antibodies by B-cells; and/or the activation of suppressor T-cells and/or memory/effector T-cells directed specifically to an antigen or antigens present in the composition or vaccine of interest. In some embodiments, these responses may serve to neutralize infectivity, and/or mediate antibody-complement, or antibody dependent cell cytotoxicity (ADCC) to provide protection to an immunized host. Such responses can be determined using standard immunoassays and neutralization assays, well known in the art. (See, *e.g.*, Montefiori *et al.*, 1988, J Clin Microbiol. 26:231-235; Dreyer *et al.*, 1999, AIDS Res Hum Retroviruses 15(17):1563-1571). The innate immune system of mammals also recognizes and responds to molecular features of pathogenic organisms and cancer cells *via* activation of Toll-like receptors and similar receptor molecules on immune cells. Upon activation of the innate immune system, various non-adaptive immune response cells are activated to, *e.g.*, produce various cytokines, lymphokines and chemokines. Cells activated by an innate immune response include immature and mature dendritic cells of, for example, the monocyte and plamsacytoid lineage (MDC, PDC), as well as gamma, delta, alpha and beta T cells and B cells and the like. Thus, the present invention also contemplates an immune response wherein the immune response involves both an innate and adaptive response.

[0060] A composition is "immunogenic" if it is capable of either: a) generating an immune response against a target antigen (*e.g.*, a viral or tumor antigen) in an individual; or b) reconstituting, boosting, or maintaining an immune response in an individual beyond what would occur if the agent or composition was not administered. An agent or composition is immunogenic if it is capable of attaining either of these criteria when administered in single or multiple doses.

[0061] "Immunomodulation," modulating an immune response" and the like refer to the modulation of the immune system in response to a stimulus and includes increasing or decreasing an immune response to a target antigen or changing an immune response from one that is predominantly a humoral immune response to one that is a more cell-mediated immune response and *vice versa*. For example, it is known in the art that decreasing the amount of antigen for immunization can change the bias of the immune system from a predominantly humoral immune response to a predominantly cellular immune response.

[0062] By "isoaccepting transfer RNA" or "iso-tRNA" is meant one or more transfer RNA molecules that differ in their anticodon nucleotide sequence but are specific for the same amino acid.

[0063] As used herein, the term "mammal" refers to any mammal including, without limitation, humans and other primates, including non-human primates such as chimpanzees and other apes and monkey species; farm animals such as cattle, sheep, pigs, goats and horses; domestic mammals such as dogs and cats; and laboratory animals including rodents such as mice, rats and guinea pigs. The term does not denote a particular age. Thus, both adult and newborn individuals are intended to be covered.

[0064] By "modulating," "modulate" and the like is meant increasing or decreasing, either directly or indirectly, the quality of a selected phenotype (*e.g.*, an immune response). In certain embodiments, "modulation" or "modulating" means that a desired/selected immune response is more efficient (*e.g.*, at least 10%, 20%, 30%, 40%, 50%, 60% or more), more rapid (*e.g.*, at least 10%, 20%, 30%, 40%, 50%, 60% or more), greater in magnitude (*e.g.*, at least 10%, 20%, 30%, 40%, 50%, 60% or more), and/or more easily induced (*e.g.*, at least 10%, 20%, 30%, 40%, 50%, 60% or more) than if the parent polynucleotide had been used under the same conditions as the synthetic polynucleotide. In other embodiments, "modulation" or "modulating" means changing an immune response from a predominantly antibody-mediated immune response as conferred by the parent polynucleotide, to a predominantly cellular immune response as conferred by the synthetic polynucleotide under the same conditions. In still other embodiments, "modulation" or "modulating" means changing an immune response from a predominantly cellular immune response as conferred by the parent polynucleotide, to a predominantly antibody-mediated immune response as conferred by the synthetic polynucleotide under the same conditions.

[0065] By "natural gene" is meant a gene that naturally encodes the protein. However, it is possible that the parent polynucleotide encodes a protein that is not naturally-occurring but has been engineered using recombinant techniques.

[0066] The term "5' non-coding region" is used herein in its broadest context to include all nucleotide sequences which are derived from the upstream region of an expressible gene, other than those sequences which encode amino acid residues which comprise the polypeptide product of the gene, wherein 5' non-coding region confers or activates or otherwise facilitates, at least in part, expression of the gene.

[0067] The term "oligonucleotide" as used herein refers to a polymer composed of a multiplicity of nucleotide units (deoxyribonucleotides or ribonucleotides, or related structural variants or synthetic analogues thereof) linked *via* phosphodiester bonds (or related structural variants or synthetic analogues thereof). Thus, while the term "oligonucleotide" typically refers to a nucleotide polymer in which the nucleotides and linkages between them are naturally occurring, it will be understood that the term also includes within its scope various analogues including, but not restricted to, peptide nucleic acids (PNAs), phosphoramidates, phosphorothioates, methyl phosphonates, 2-O-methyl ribonucleic acids, and the like. The exact size of the molecule may vary depending on the particular application. An oligonucleotide is typically rather short in length, generally from about 10 to 30 nucleotides, but the term can refer to molecules of any length, although the term "polynucleotide" or "nucleic acid" is typically used for large oligonucleotides.

[0068] The terms "operably connected," "operably linked" and the like as used herein refer to an arrangement of elements wherein the components so described are configured so as to perform their usual function. Thus, a given promoter operably linked to a coding sequence is capable of effecting the expression of the coding sequence when the proper enzymes are present. The promoter need not be contiguous with the coding sequence, so long as it functions to direct the expression thereof. Thus, for example, intervening untranslated yet transcribed sequences can be present between the promoter sequence and the coding sequence and the promoter sequence can still be considered "operably linked" to the coding sequence. Terms such as "operably connected," therefore, include placing a structural gene under the regulatory control of a promoter, which then controls the transcription and optionally translation of the gene. In the construction of heterologous promoter/structural gene combinations, it is generally preferred to position the genetic sequence or promoter at a distance from the gene transcription start site that is approximately the same as the distance

between that genetic sequence or promoter and the gene it controls in its natural setting; *i.e.* the gene from which the genetic sequence or promoter is derived. As is known in the art, some variation in this distance can be accommodated without loss of function. Similarly, the preferred positioning of a regulatory sequence element with respect to a heterologous gene to be placed under its control is defined by the positioning of the element in its natural setting; *i.e.*, the genes from which it is derived.

[0069] By "pharmaceutically-acceptable carrier" is meant a solid or liquid filler, diluent or encapsulating substance that may be safely used in topical or systemic administration.

[0070] The term "phenotype" means any one or more detectable physical or functional characteristics, properties, attributes or traits of an organism, tissue, or cell, or class of organisms, tissues or cells, which generally result from the interaction between the genetic makeup (*i.e.*, genotype) of the organism, tissue, or cell, or the class of organisms, tissues or cells and the environment.

[0071] By "phenotypic preference" is meant the preference with which an organism uses a codon to produce a selected phenotype. This preference can be evidenced, for example, by the quality of a selected phenotype that is producible by a polynucleotide that comprises the codon in an open reading frame which codes for a polypeptide that produces the selected phenotype. In certain embodiment, the preference of usage is independent of the route by which the polynucleotide is introduced into the organism. However, in other embodiments, the preference of usage is dependent on the route of introduction of the polynucleotide into the organism.

[0072] The term "polynucleotide" or "nucleic acid" as used herein designates mRNA, RNA, cRNA, cDNA or DNA. The term typically refers to oligonucleotides greater than 30 nucleotides in length.

[0073] "Polypeptide," "peptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues and to variants and synthetic analogues of the same. Thus, these terms apply to amino acid polymers in which one or more amino acid residues is a synthetic non-naturally occurring amino acid, such as a chemical analogue of a corresponding naturally occurring amino acid, as well as to naturally-occurring amino acid polymers. As used herein, the terms "polypeptide," "peptide" and "protein" are not limited to a minimum length of the product. Thus, peptides, oligopeptides, dimers, multimers, and the like, are

included within the definition. Both full-length proteins and fragments thereof are encompassed by the definition. The terms also include post expression modifications of a polypeptide, for example, glycosylation, acetylation, phosphorylation and the like. In some embodiments, a "polypeptide" refers to a protein which includes modifications, such as deletions, additions and substitutions (generally conservative in nature), to the native sequence, so long as the protein maintains the desired activity. These modifications may be deliberate, as through site-directed mutagenesis, or may be accidental, such as through mutations of hosts which produce the proteins or errors due to PCR amplification.

[0074] The terms "polypeptide variant," and "variant" refer to polypeptides that vary from a reference polypeptide by the addition, deletion or substitution (generally conservative in nature) of at least one amino acid residue. Typically, variants retain a desired activity of the reference polypeptide, such as antigenic activity in inducing an immune response against a target antigen. In general, variant polypeptides are "substantially similar" or substantially identical" to the reference polypeptide, *e.g.*, amino acid sequence identity or similarity of more than 50%, generally more than 60%-70%, even more particularly 80%-85% or more, such as at least 90%-95% or more, when the two sequences are aligned. Often, the variants will include the same number of amino acids but will include substitutions, as explained herein.

[0075] By "primer" is meant an oligonucleotide which, when paired with a strand of DNA, is capable of initiating the synthesis of a primer extension product in the presence of a suitable polymerizing agent. The primer is preferably single-stranded for maximum efficiency in amplification but may alternatively be double-stranded. A primer must be sufficiently long to prime the synthesis of extension products in the presence of the polymerization agent. The length of the primer depends on many factors, including application, temperature to be employed, template reaction conditions, other reagents, and source of primers. For example, depending on the complexity of the target sequence, the oligonucleotide primer typically contains 15 to 35 or more nucleotides, although it may contain fewer nucleotides. Primers can be large polynucleotides, such as from about 200 nucleotides to several kilobases or more. Primers may be selected to be "substantially complementary" to the sequence on the template to which it is designed to hybridize and serve as a site for the initiation of synthesis. By "substantially complementary", it is meant that the primer is sufficiently complementary to hybridize with a target nucleotide sequence. Preferably, the primer contains no mismatches with the template to which it is designed to

hybridize but this is not essential. For example, non-complementary nucleotides may be attached to the 5' end of the primer, with the remainder of the primer sequence being complementary to the template. Alternatively, non-complementary nucleotides or a stretch of non-complementary nucleotides can be interspersed into a primer, provided that the primer sequence has sufficient complementarity with the sequence of the template to hybridize therewith and thereby form a template for synthesis of the extension product of the primer.

[0076] Reference herein to a "promoter" is to be taken in its broadest context and includes the transcriptional regulatory sequences of a classical genomic gene, including the TATA box which is required for accurate transcription initiation, with or without a CCAAT box sequence and additional regulatory elements (*i.e.* upstream activating sequences, enhancers and silencers) which alter gene expression in response to developmental and/or environmental stimuli, or in a tissue-specific or cell-type-specific manner. A promoter is usually, but not necessarily, positioned upstream or 5', of a structural gene, the expression of which it regulates. Furthermore, the regulatory elements comprising a promoter are usually positioned within 2 kb of the start site of transcription of the gene. Preferred promoters according to the invention may contain additional copies of one or more specific regulatory elements to further enhance expression in a cell, and/or to alter the timing of expression of a structural gene to which it is operably connected.

[0077] The term "quality" is used herein in its broadest sense and includes a measure, strength, intensity, degree or grade of a phenotype, *e.g.*, a superior or inferior immune response.

[0078] The term "sequence identity" as used herein refers to the extent that sequences are identical on a nucleotide-by-nucleotide basis or an amino acid-by-amino acid basis over a window of comparison. Thus, a "percentage of sequence identity" is calculated by comparing two optimally aligned sequences over the window of comparison, determining the number of positions at which the identical nucleic acid base (*e.g.*, A, T, C, G, I) or the identical amino acid residue (*e.g.*, Ala, Pro, Ser, Thr, Gly, Val, Leu, Ile, Phe, Tyr, Trp, Lys, Arg, His, Asp, Glu, Asn, Gln, Cys and Met) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison (*i.e.*, the window size), and multiplying the result by 100 to yield the percentage of sequence identity. For the purposes of the present invention, "sequence identity" will be understood to mean the "match percentage" calculated by the DNASIS computer program (Version 2.5 for windows; available from Hitachi Software

engineering Co., Ltd., South San Francisco, California, USA) using standard defaults as used in the reference manual accompanying the software.

[0079] "Similarity" refers to the percentage number of amino acids that are identical or constitute conservative substitutions as defined in Table 10. Similarity may be determined using sequence comparison programs such as GAP (Deveraux *et al.* 1984, Nucleic Acids Research 12, 387-395). In this way, sequences of a similar or substantially different length to those cited herein might be compared by insertion of gaps into the alignment, such gaps being determined, for example, by the comparison algorithm used by GAP.

[0080] Terms used to describe sequence relationships between two or more polynucleotides or polypeptides include "reference sequence", "comparison window", "sequence identity", "percentage of sequence identity" and "substantial identity". A "reference sequence" is at least 12 but frequently 15 to 18 and often at least 25 monomer units, inclusive of nucleotides and amino acid residues, in length. Because two polynucleotides may each comprise (1) a sequence (*i.e.*, only a portion of the complete polynucleotide sequence) that is similar between the two polynucleotides, and (2) a sequence that is divergent between the two polynucleotides, sequence comparisons between two (or more) polynucleotides are typically performed by comparing sequences of the two polynucleotides over a "comparison window" to identify and compare local regions of sequence similarity. A "comparison window" refers to a conceptual segment of at least 6 contiguous positions, usually about 50 to about 100, more usually about 100 to about 150 in which a sequence is compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. The comparison window may comprise additions or deletions (*i.e.*, gaps) of about 20% or less as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. Optimal alignment of sequences for aligning a comparison window may be conducted by computerized implementations of algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package Release 7.0, Genetics Computer Group, 575 Science Drive Madison, WI, USA) or by inspection and the best alignment (*i.e.*, resulting in the highest percentage homology over the comparison window) generated by any of the various methods selected. Reference also may be made to the BLAST family of programs as for example disclosed by Altschul *et al.*, 1997, Nucl. Acids Res. 25:3389. A detailed discussion of sequence analysis can be found in Unit 19.3 of Ausubel *et al.*, "Current Protocols in Molecular Biology", John Wiley & Sons Inc, 1994-1998, Chapter 15.

[0081] The term "synthetic polynucleotide" as used herein refers to a polynucleotide that is formed by recombinant or synthetic techniques and typically includes polynucleotides that are not normally found in nature.

[0082] The term "synonymous codon" as used herein refers to a codon having a different nucleotide sequence than another codon but encoding the same amino acid as that other codon.

[0083] By "treatment," "treat," "treated" and the like is meant to include both therapeutic and prophylactic treatment.

[0084] By "vector" is meant a nucleic acid molecule, preferably a DNA molecule derived, for example, from a plasmid, bacteriophage, or plant virus, into which a nucleic acid sequence may be inserted or cloned. A vector preferably contains one or more unique restriction sites and may be capable of autonomous replication in a defined host cell including a target cell or tissue or a progenitor cell or tissue thereof, or be integrable with the genome of the defined host such that the cloned sequence is reproducible. Accordingly, the vector may be an autonomously replicating vector, *i.e.*, a vector that exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, *e.g.*, a linear or closed circular plasmid, an extrachromosomal element, a minichromosome, or an artificial chromosome. The vector may contain any means for assuring self-replication. Alternatively, the vector may be one which, when introduced into the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. A vector system may comprise a single vector or plasmid, two or more vectors or plasmids, which together contain the total DNA to be introduced into the genome of the host cell, or a transposon. The choice of the vector will typically depend on the compatibility of the vector with the host cell into which the vector is to be introduced. The vector may also include a selection marker such as an antibiotic resistance gene that can be used for selection of suitable transformants. Examples of such resistance genes are well known to those of skill in the art.

2. *Abbreviations*

[0085] The following abbreviations are used throughout the application:

nt = nucleotide
nts = nucleotides
aa = amino acid(s)
kb = kilobase(s) or kilobase pair(s)
kDa = kilodalton(s)
d = day
h = hour
s = seconds

3. *Immune response preference ranking of codons in mammals*

[0086] The present invention provides for the first time an immune response preference ranking of individual synonymous codons in mammals. This ranking was determined using a construct system that comprises a series of reporter constructs each comprising a different coding sequence for an antigenic polypeptide (*e.g.*, a papillomavirus E7 polypeptide), wherein the coding sequence of individual constructs is distinguished from a parent coding sequence that encodes the antigenic polypeptide by the substitution of a single species of iso-accepting codon for each other species of iso-accepting codon that is present in the parent coding sequence. Accordingly, the coding sequence of individual synthetic constructs uses the same iso-accepting codon to encode most instances and preferably every instance of a particular amino acid residue (*e.g.*, Ala^{GCT} for all alanines) in the antigenic polypeptide and individual synthetic constructs differ in the species of iso-accepting codon used to encode a particular amino acid residue across the polypeptide sequence. As used herein, the species of iso-accepting codon that is used to encode a particular amino acid residue in the antigenic polypeptide is referred to as a "standardized codon". An illustrative synthetic construct system is described in Example 1, which covers the entire set of synonymous codons that code for amino acids.

[0087] Test mammals (*e.g.*, mice) were immunized with the synthetic construct system in which individual mammals were immunized with a different synthetic construct and the host immune response (*e.g.*, a humoral immune response or a cellular immune response) to the antigenic polypeptide was determined for each construct. In accordance with the

present invention, the strength of immune response obtained from individual synthetic constructs provides a direct correlation to the immune preference of a corresponding standardized codon in a test mammal. Accordingly, the stronger the immune response produced from a given construct in a test mammal, the higher the immune preference will be of the corresponding standardized codon.

[0088] Comparison of the immune response preferences so determined with the translational efficiencies derived from codon usage frequency values for mammalian cells in general as determined by Seed (see U.S. Patent Serial Nos 5,786,464 and 5,795,737) reveals several differences in the ranking of codons. For convenience, these differences are highlighted in TABLE 9, in which Seed 'preferred' codons are highlighted with a blue background, Seed 'less preferred' codons are highlighted with a green background, and Seed 'non preferred' codons are highlighted with a grey background.

TABLE 9

aa	Preferential codon usage as predicted by Seed for mammalian cells in general	Experimentally determined codon immune response preferences in test mammals
Ala	GCC >> (GCG, GCT, GCA)	GCT > GCC > (GCA GCG)
Arg	CGC >> (CGA, CGT, AGA, AGG, CGG)	(CGA, CGC, CGT, AGA) > (AGG, CGG)
Asn	AAC >> AAT	AAC > AAT
Asp	GAC >> GAT	GAC > GAT
Cys	TGC >> TGT	TGC > TGT
Glu	(GAA, GAG)	GAA > GAG
Gln	CAG >> CAA	CAA = CAG
Gly	GGC > GGG > (GGT, GGA)	GGA > (GGG, GGT, GGC)
His	CAC >> CAT	CAC = CAT
Ile	ATC > ATT > ATA	ATC >> ATT > ATA
Leu	CTG > CTC > (TTA,CTA, CTT, TTG)	(CTG, CTC) > (CTA, CTT) >> TTG > TTA
Lys	AAG >> AAA	AAG = AAA

aa	Preferential codon usage as predicted by Seed for mammalian cells in general	Experimentally determined codon immune response preferences in test mammals
Phe	TTC >> TTT	TTT > TTC
Pro	CCC >> (CCG, CCA, CCT)	CCC > CCT >> (CCA, CCG)
Ser	AGC > TCC > (TCG, AGT, TCA, TCT)	TCG >> (TCT, TCA, TCC) >> (AGC, AGT)
Thr	ACC >> (ACG, ACA, ACT)	ACG > ACC >> ACA > ACT
Tyr	TAC >> TAT	TAC > TAT
Val	GTG > GTC > (GTA, GTT)	(GTG, GTC) > GTT > GTA

[0089] As will be apparent from the above table:

[0090] (i) several codons deemed by Seed to have a higher codon usage ranking in mammalian cells than at least one other synonymous codon have in fact a lower immune response preference ranking than the or each other synonymous codon (*e.g.*, Ala^{GCC} has a higher codon usage ranking but lower immune response preference ranking than Ala^{GCT}; Gly^{GGC} has a higher codon usage ranking but lower immune response preference ranking than Gly^{GGA}; Phe^{TTC} has a higher codon usage ranking but lower immune response preference ranking than Phe^{TTT}; Ser^{AGC} has a higher codon usage ranking but lower immune response preference ranking than any one of Ser^{TCG}, Ser^{TCT}, Ser^{TCG}, Ser^{TCA} and Ser^{TCC}; and Thr^{ACC} has a higher codon usage ranking but lower immune response preference ranking than Thr^{ACG});

[0091] (ii) several codons deemed by Seed to have a lower codon usage ranking in mammalian cells than at least one other synonymous codon have in fact a higher immune response preference ranking than the or each other synonymous codon (*e.g.*, Ala^{GCT} has a lower codon usage ranking but higher immune response preference ranking than Ala^{GCC}; Gly^{GGA} has a lower codon usage ranking but higher immune response preference ranking than Gly^{GGC} or Gly^{GGG}; Phe^{TTT} has a lower codon usage ranking but higher immune response preference ranking than Phe^{TTC}; Ser^{TCG} has a lower codon usage ranking but higher immune response preference ranking than Ser^{AGC} or Ser^{TCC}; Ser^{TCT} and Ser^{TCA} have a lower codon usage ranking but higher immune response preference ranking than Ser^{AGC}; and Thr^{ACG} has a lower codon usage ranking but higher immune response preference ranking than Thr^{ACC});

[0092] (iii) several codons deemed by Seed to have a higher codon usage ranking in mammalian cells than another synonymous codon have in fact the same immune response preference ranking as the other synonymous codon (*e.g.*, Gln^{CAG} has a higher codon usage ranking than, but the same immune response preference ranking as, Gln^{CAA}; His^{CAC} has a higher codon usage ranking than, but the same immune response preference ranking as, His^{CAT}; Leu^{CTG} has a higher codon usage ranking than, but the same immune response preference ranking as Leu^{CTC}; Lys^{AAG} has a higher codon usage ranking than, but the same immune response preference ranking as, Lys^{AAA}; Val^{GTG} has a higher codon usage ranking than, but the same immune response preference ranking as, Val^{GTC}); and

[0093] (iv) several codons deemed by Seed to have the same codon usage ranking in mammalian cells as at least one other synonymous codon have in fact a different immune response preference ranking than the or each other synonymous codon (*e.g.*, Ala^{GCT} has the same codon usage ranking as, but a higher immune response preference ranking than, Ala^{GCA} and Ala^{GCG}; Arg^{CGA}, Arg^{CGT} and Arg^{AGA} have the same codon usage ranking as, but a higher immune response preference ranking than, Arg^{AGG} and Arg^{CGG}; Glu^{GAA} has the same codon usage ranking as, but a higher immune response preference ranking than, Glu^{GAG}; Gly^{GGA} has the same codon usage ranking as, but a higher immune response preference ranking than, Gly^{GGT}; Leu^{CTA} and Leu^{CTT} have the same codon usage ranking as, but a higher immune response preference ranking than, Leu^{TTG} and Leu^{TTA}; Pro^{CCT} has the same codon usage ranking as, but a higher immune response preference ranking than, Pro^{CCA} or Pro^{CCG}; Ser^{TCG} has the same codon usage ranking as, but a higher immune response preference ranking than, any one of Ser^{TCT}, Ser^{TCA} and Ser^{AGT}; Ser^{TCT} and Ser^{TCA} have the same codon usage ranking as, but a higher immune response preference ranking than, Ser^{AGT}; Thr^{ACG} has the same codon usage ranking as, but a higher immune response preference ranking than, any one of Thr^{ACA} and Thr^{ACT}; Thr^{ACG} has the same codon usage ranking as, but a higher immune response preference ranking than, Thr^{ACT}; Val^{GTT} has the same codon usage ranking as, but a higher immune response preference ranking than, Val^{GTA}).

[0094] Accordingly, the present invention enables for the first time the modulation of an immune response to a target antigen in a mammal from a polynucleotide that encodes a polypeptide that corresponds to at least a portion of the target antigen by replacing at least one codon of the polynucleotide with a synonymous codon that has a higher or lower preference for producing an immune response than the codon it replaces. In some embodiments, therefore, the present invention embraces methods of constructing a synthetic polynucleotide

from which a polypeptide is producible to confer an enhanced or stronger immune response than one conferred by a parent polynucleotide that encodes the same polypeptide. These methods generally comprise selecting from TABLE 1 a codon (often referred to herein arbitrarily as a "first codon") of the parent polynucleotide for replacement with a synonymous codon, wherein the synonymous codon is selected on the basis that it exhibits a higher immune response preference than the first codon and replacing the first codon with the synonymous codon to construct the synthetic polynucleotide. Illustrative selections of the first and synonymous codons are made according to TABLE 2.

[0095] In some embodiments, the selection of the first and synonymous codons is made according to TABLE 3, which is the same as TABLE 2 with the exception that it excludes selections based on codon usage rankings as disclosed by Seed. In illustrative examples of this type, the selection of a second codon (and subsequent codons if desired) for replacement with a synonymous codon is made according to TABLE 4.

[0096] Where synonymous codons are classified into three ranks ('high', 'intermediate' and 'low' ranks) based on their immune response preference ranking (*e.g.*, the synonymous codons for Ala, Ile, Leu, Pro, Ser, Thr and Val), it is preferred that the synonymous codon that is selected is a high rank codon when the first codon is a low rank codon. However, this is not essential and the synonymous codon can be selected from intermediate rank codons. In the case of two or more synonymous codons having similar immune response preferences, it will be appreciated that any one of these codons can be used to replace the first codon.

[0097] In other embodiments, the invention provides methods of constructing a synthetic polynucleotide from which a polypeptide is producible to confer a reduced or weaker immune response than one conferred by a parent polynucleotide that encodes the same polypeptide. These methods generally comprise selecting from TABLE 1 a first codon of the parent polynucleotide for replacement with a synonymous codon, wherein the synonymous codon is selected on the basis that it exhibits a lower immune response preference than the first codon and replacing the first codon with the synonymous codon to construct the synthetic polynucleotide. Illustrative selections of the first and synonymous codons are made according to TABLE 5.

[0098] In some embodiments, the selection of the first and synonymous codons is made according to TABLE 6, which is the same as TABLE 5 with the exception that it excludes selections based on codon usage rankings as disclosed by Seed. In illustrative

examples of this type, the selection of a second codon (and subsequent codons if desired) for replacement with a synonymous codon is made according to TABLE 7.

[0099] Where synonymous codons are classified into the three ranks noted above, it is preferred that the synonymous codon that is selected is a low rank codon when the first codon is a high rank codon but this is not essential and thus the synonymous codon can be selected from intermediate rank codons if desired.

[0100] Generally, the difference in strength of the immune response produced in the mammal from the synthetic polynucleotide as compared to that produced from the parent polynucleotide depends on the number of first/second codons that are replaced by synonymous codons, and on the difference in immune response preference ranking between the first/second codons and the synonymous codons. Put another way, the fewer such replacements, and/or the smaller the difference in immune response preference ranking between the synonymous and first/codons codons, the smaller the difference will be in the immune response produced by the synthetic polynucleotide and the one produced by the parent polynucleotide. Conversely, the more such replacements, and/or the greater the difference in immune response preference ranking between the synonymous and first/second codons, the greater the difference will be in the immune response produced by the synthetic polynucleotide and the one produced by the parent polynucleotide.

[0101] It is preferable but not necessary to replace all the codons of the parent polynucleotide with synonymous codons having different (*e.g.*, higher or lower) immune response preference rankings than the first/second codons. Changes in the conferred immune response can be accomplished even with partial replacement. Generally, the replacement step affects at least about 5%, 10%, 15%, 20%, 25%, 30%, usually at least about 35%, 40%, 50%, and typically at least about 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, 99% or more of the first/second codons of the parent polynucleotide. In embodiments in which a stronger or enhanced immune response is required, it is generally desirable to replace some, preferably most and more preferably all, low rank codons in a parent polynucleotide with synonymous codons that are intermediate, or preferably high rank codons. Typically, replacement of low with intermediate or high rank codons will result in an increase in the strength of immune response from the synthetic polynucleotide so constructed, as compared to the one produced from the parent polynucleotide under the same conditions. However, it is often desirable to replace some, preferably most and more preferably all, intermediate rank codons in the parent

polynucleotide with high rank codons, if stronger or more enhanced immune responses are desired.

[0102] By contrast, in some embodiments in which a weaker or reduced immune response is required, it is generally desirable to replace some, preferably most and more preferably all, high rank codons in a parent polynucleotide with synonymous codons that are intermediate, or preferably low rank codons. Typically, replacement of high with intermediate or low rank codons will result in a substantial decrease in the strength of immune response from the synthetic polynucleotide so constructed, as compared to the one produced from the parent polynucleotide under the same condition. In specific embodiments in which it is desired to confer a weaker or more reduced immune response, it is generally desirable to replace some, preferably most and more preferably all, intermediate rank codons in the parent polynucleotide with low rank codons.

[0103] In illustrative examples requiring a stronger or enhanced immune response, the number of, and difference in immune response preference ranking between, the first/second codons and the synonymous codons are selected such that the immune response conferred by the synthetic polynucleotide is at least about 110%, 150%, 200%, 250%, 300%, 350%, 400%, 450%, 500%, 600%, 700%, 800%, 900%, 1000%, or more, of the immune response conferred by the parent polynucleotide under the same conditions. Conversely, in some embodiments requiring a lower or weaker immune response, the number of, and difference in phenotypic preference ranking between, the first/second codons and the synonymous codons are selected such that the immune response conferred by the synthetic polynucleotide is no more than about 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5%, or less of the immune response conferred by the parent polynucleotide under the same conditions.

4. Modulating immune responses in mammals by expression of isoaccepting transfer RNA-encoding polynucleotides

[0104] It is possible to take advantage of the immune response preference rankings of codons discussed in Section 3 to modulate an immune response to a target antigen by changing the level of iso-tRNAs in the cell population which is the target of the immunization. Accordingly, the invention also features methods of enhancing the quality of an immune response to a target antigen in a mammal, wherein the response is conferred by the expression of a first polynucleotide that encodes a polypeptide corresponding to at least a

portion of the target antigen. These methods generally comprise: introducing into the mammal a first nucleic acid construct comprising the first polynucleotide in operable connection with a regulatory polynucleotide. A second nucleic acid construct is then introduced into the mammal, which comprises a second polynucleotide that is operably connected to a regulatory polynucleotide and that encodes an iso-tRNA corresponding to a low immune preference codon of the first polynucleotide.

[0105] In practice, therefore, an iso-tRNA is introduced into the mammal by the second nucleic acid construct when the iso-tRNA corresponds to a low immune response preference codon in the first polynucleotide, which are suitably selected from the group consisting of Ala^{GCA}, Ala^{GCG}, Ala^{GCC}, Arg^{AGG}, Arg^{CGG}, Asn^{AAT}, Asp^{GAT}, Cys^{TGT}, Glu^{GAG}, Gly^{GGG}, Gly^{GGT}, Gly^{GGC}, Ile^{ATA}, Ile^{ATT}, Leu^{TTG}, Leu^{TTA}, Leu^{CTA}, Leu^{CTT}, Phe^{TTC}, Pro^{CCA}, Pro^{CCG}, Pro^{CCT}, Ser^{AGC}, Ser^{AGT}, Ser^{TCT}, Ser^{TCA}, Ser^{TCC}, Thr^{ACA}, Thr^{ACT}, Tyr^{TAT}, Val^{GTA} and Val^{GTT}. In specific embodiments, the supplied iso-tRNAs are specific for codons that have 'low' immune response preference codons, which may be selected from the group consisting of Ala^{GCA}, Ala^{GCG}, Arg^{AGG}, Arg^{CGG}, Asn^{AAT}, Asp^{GAT}, Cys^{TGT}, Glu^{GAG}, Gly^{GGG}, Gly^{GGT}, Gly^{GGC}, Ile^{ATA}, Leu^{TTG}, Leu^{TTA}, Phe^{TTC}, Pro^{CCA}, Pro^{CCG}, Ser^{AGC}, Ser^{AGT}, Thr^{ACT}, Tyr^{TAT} and Val^{GTA}. The first construct (*i.e.*, antigen-expressing construct) and the second construct (*i.e.*, the iso-tRNA-expressing construct) may be introduced simultaneously or sequentially (in either order) and may be introduced at the same or different sites. In some embodiments, the first and second constructs are contained in separate vectors. In other embodiments, they are contained in a single vector. If desired, two or more second constructs may be introduced each expressing a different iso-tRNA corresponding to a low preference codon of the first polynucleotide. The first and second nucleic acid constructs may be constructed and administered concurrently or contemporaneously to a mammal according to any suitable method, illustrative examples of which are discussed below for the chimeric constructs of the invention.

[0106] In some embodiments, a plurality of different iso-tRNA-expressing constructs (*e.g.*, 2, 3, 4, 5, 6,7 ,8 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 or more) are administered concurrently or contemporaneously with the antigen-expressing construct, wherein individual iso-tRNA-expressing constructs express a different iso-tRNA than other iso-tRNA-expressing constructs.

5. Antigens

[0107] Target antigens useful in the present invention are typically proteinaceous molecules, representative examples of which include polypeptides and peptides. Target antigens may be selected from endogenous antigens produced by a host or exogenous antigens that are foreign to the host. Suitable endogenous antigens include, but are not restricted to, cancer or tumor antigens. Non-limiting examples of cancer or tumor antigens include antigens from a cancer or tumor selected from ABL1 proto-oncogene, AIDS related cancers, acoustic neuroma, acute lymphocytic leukemia, acute myeloid leukemia, adenocystic carcinoma, adrenocortical cancer, agnogenic myeloid metaplasia, alopecia, alveolar soft-part sarcoma, anal cancer, angiosarcoma, aplastic anemia, astrocytoma, ataxia-telangiectasia, basal cell carcinoma (skin), bladder cancer, bone cancers, bowel cancer, brain stem glioma, brain and CNS tumors, breast cancer, CNS tumors, carcinoid tumors, cervical cancer, childhood brain tumors, childhood cancer, childhood leukemia, childhood soft tissue sarcoma, chondrosarcoma, choriocarcinoma, chronic lymphocytic leukemia, chronic myeloid leukemia, colorectal cancers, cutaneous T-cell lymphoma, dermatofibrosarcoma protuberans, desmoplastic small round cell tumor, ductal carcinoma, endocrine cancers, endometrial cancer, ependymoma, oesophageal cancer, Ewing's Sarcoma, Extra-Hepatic Bile Duct Cancer, Eye Cancer, Eye: Melanoma, Retinoblastoma, Fallopian Tube cancer, Fanconi anemia, fibrosarcoma, gall bladder cancer, gastric cancer, gastrointestinal cancers, gastrointestinal-carcinoid-tumor, genitourinary cancers, germ cell tumors, gestational-trophoblastic-disease, glioma, gynecological cancers, haematological malignancies, hairy cell leukemia, head and neck cancer, hepatocellular cancer, hereditary breast cancer, histiocytosis, Hodgkin's disease, human papillomavirus, hydatidiform mole, hypercalcemia, hypopharynx cancer, intraocular melanoma, islet cell cancer, Kaposi's sarcoma, kidney cancer, Langerhans cell histiocytosis, laryngeal cancer, leiomyosarcoma, leukemia, Li-Fraumeni syndrome, lip cancer, liposarcoma, liver cancer, lung cancer, lymphedema, lymphoma, Hodgkin's lymphoma, non-Hodgkin's lymphoma, male breast cancer, malignant-rhabdoid tumor of kidney, medulloblastoma, melanoma, Merkel cell cancer, mesothelioma, metastatic cancer, mouth cancer, multiple endocrine neoplasia, mycosis fungoides, myelodysplastic syndromes, myeloma, myeloproliferative disorders, nasal cancer, nasopharyngeal cancer, nephroblastoma, neuroblastoma, neurofibromatosis, Nijmegen breakage syndrome, non-melanoma skin cancer, non-small-cell-lung-cancer (NSCLC), ocular cancers, esophageal cancer, oral cavity cancer, oropharynx cancer, osteosarcoma, ostomy ovarian cancer, pancreas cancer, paranasal cancer, parathyroid cancer, parotid gland cancer, penile cancer, peripheral-neuroectodermal tumours,

pituitary cancer, polycythemia vera, prostate cancer, rare cancers and associated disorders, renal cell carcinoma, retinoblastoma, rhabdomyosarcoma, Rothmund-Thomson syndrome, salivary gland cancer, sarcoma, schwannoma, Sezary syndrome, skin cancer, small cell lung cancer (SCLC), small intestine cancer, soft tissue sarcoma, spinal cord tumors, squamous-cell-carcinoma-(skin), stomach cancer, synovial sarcoma, testicular cancer, thymus cancer, thyroid cancer, transitional-cell-cancer-(bladder), transitional-cell-cancer-(renal-pelvis-/-ureter), trophoblastic cancer, urethral cancer, urinary system cancer, uroplakins, uterine sarcoma, uterus cancer, vaginal cancer, vulva cancer, Waldenstroms macroglobulinemia, Wilms' tumor. In certain embodiments, the cancer or tumor relates to melanoma. Illustrative examples of melanoma-related antigens include melanocyte differentiation antigen (*e.g.*, gp100, MART, Melan-A/MART-1, TRP-1, Tyros, TRP2, MC1R, MUC1F, MUC1R or a combination thereof) and melanoma-specific antigens (*e.g.*, BAGE, GAGE-1, gp100In4, MAGE-1 (*e.g.*, GenBank Accession No. X54156 and AA494311), MAGE-3, MAGE4, PRAME, TRP2IN2, NYNSO1a, NYNSO1b, LAGE1, p97 melanoma antigen (*e.g.*, GenBank Accession No. M12154) p5 protein, gp75, oncofetal antigen, GM2 and GD2 gangliosides, cdc27, p21ras, $gp100^{Pmel117}$ or a combination thereof. Other tumour-specific antigens include, but are not limited to: etv6, aml1, cyclophilin b (acute lymphoblastic leukemia); Ig-idiotype (B cell lymphoma); E-cadherin, α-catenin, β-catenin, γ -catenin, p120ctn (glioma); p21ras (bladder cancer); p21ras (biliary cancer); MUC family, HER2/neu, c-erbB-2 (breast cancer); p53, p21ras (cervical carcinoma); p21ras, HER2/neu, c-erbB-2, MUC family, Cripto-1protein, Pim-1 protein (colon carcinoma); Colorectal associated antigen (CRC)-CO17-1A/GA733, APC (colorectal cancer); carcinoembryonic antigen (CEA) (colorectal cancer; choriocarcinoma); cyclophilin b (epithelial cell cancer); HER2/neu, c-erbB-2, ga733 glycoprotein (gastric cancer); α-fetoprotein (hepatocellular cancer); Imp-1, EBNA-1 (Hodgkin's lymphoma); CEA, MAGE-3, NY-ESO-1 (lung cancer); cyclophilin b (lymphoid cell-derived leukemia); MUC family, p21ras (myeloma); HER2/neu, c-erbB-2 (non-small cell lung carcinoma); Imp-1, EBNA-1 (nasopharyngeal cancer); MUC family, HER2/neu, c-erbB-2, MAGE-A4, NY-ESO-1 (ovarian cancer); Prostate Specific Antigen (PSA) and its antigenic epitopes PSA-1, PSA-2, and PSA-3, PSMA, HER2/neu, c-erbB-2, ga733 glycoprotein (prostate cancer); HER2/neu, c-erbB-2 (renal cancer); viral products such as human papillomavirus proteins (squamous cell cancers of the cervix and esophagus); NY-ESO-1 (testicular cancer); and HTLV-1 epitopes (T cell leukemia).

[0108] Foreign or exogenous antigens are suitably selected from antigens of pathogenic organisms. Exemplary pathogenic organisms include, but are not limited to, viruses, bacteria, fungi parasites, algae and protozoa and amoebae. Illustrative viruses include viruses responsible for diseases including, but not limited to, measles, mumps, rubella, poliomyelitis, hepatitis A, B (*e.g.*, GenBank Accession No. E02707), and C (*e.g.*, GenBank Accession No. E06890), as well as other hepatitis viruses, influenza, adenovirus (*e.g.*, types 4 and 7), rabies (*e.g.*, GenBank Accession No. M34678), yellow fever, Epstein-Barr virus and other herpesviruses such as papillomavirus, Ebola virus, influenza virus, Japanese encephalitis (*e.g.*, GenBank Accession No. E07883), dengue (*e.g.*, GenBank Accession No. M24444), hantavirus, Sendai virus, respiratory syncytial virus, orthomyxoviruses, vesicular stomatitis virus, visna virus, cytomegalovirus and human immunodeficiency virus (HIV) (*e.g.*, GenBank Accession No. U18552). Any suitable antigen derived from such viruses are useful in the practice of the present invention. For example, illustrative retroviral antigens derived from HIV include, but are not limited to, antigens such as gene products of the gag, pol, and env genes, the Nef protein, reverse transcriptase, and other HIV components. Illustrative examples of hepatitis viral antigens include, but are not limited to, antigens such as the S, M, and L proteins of hepatitis B virus, the pre-S antigen of hepatitis B virus, and other hepatitis, *e.g.*, hepatitis A, B, and C, viral components such as hepatitis C viral RNA. Illustrative examples of influenza viral antigens include; but are not limited to, antigens such as hemagglutinin and neuraminidase and other influenza viral components. Illustrative examples of measles viral antigens include, but are not limited to, antigens such as the measles virus fusion protein and other measles virus components. Illustrative examples of rubella viral antigens include, but are not limited to, antigens such as proteins El and E2 and other rubella virus components; rotaviral antigens such as VP7sc and other rotaviral components. Illustrative examples of cytomegaloviral antigens include, but are not limited to, antigens such as envelope glycoprotein B and other cytomegaloviral antigen components. Non-limiting examples of respiratory syncytial viral antigens include antigens such as the RSV fusion protein, the M2 protein and other respiratory syncytial viral antigen components. Illustrative examples of herpes simplex viral antigens include, but are not limited to, antigens such as immediate early proteins, glycoprotein D, and other herpes simplex viral antigen components. Non-limiting examples of varicella zoster viral antigens include antigens such as 9PI, gpII, and other varicella zoster viral antigen components. Non-limiting examples of Japanese encephalitis viral antigens include antigens such as proteins E, M-E, M-E-NS 1, NS 1, NS 1-NS2A, 80%E, and other Japanese encephalitis viral antigen

components. Representative examples of rabies viral antigens include, but are not limited to, antigens such as rabies glycoprotein, rabies nucleoprotein and other rabies viral antigen components. Illustrative examples of papillomavirus antigens include, but are not limited to, the L1 and L2 capsid proteins as well as the E6/E7 antigens associated with cervical cancers, See Fundamental Virology, Second Edition, eds. Fields, B.N. and Knipe, D.M., 1991, Raven Press, New York, for additional examples of viral antigens.

[0109] Illustrative examples of fungi include *Acremonium* spp., *Aspergillus* spp., *Basidiobolus* spp., *Bipolaris* spp., *Blastomyces dermatidis*, *Candida* spp., *Cladophialophora carrionii*, *Coccidioides immitis*, *Conidiobolus* spp., *Cryptococcus* spp., *Curvularia* spp., *Epidermophyton* spp., *Exophiala jeanselmei*, *Exserohilum* spp., *Fonsecaea compacta*, *Fonsecaea pedrosoi*, *Fusarium oxysporum*, *Fusarium solani*, *Geotrichum candidum*, *Histoplasma capsulatum* var. capsulatum, *Histoplasma capsulatum* var. duboisii, *Hortaea werneckii*, *Lacazia loboi*, *Lasiodiplodia theobromae*, *Leptosphaeria senegalensis*, *Madurella grisea*, *Madurella mycetomatis*, *Malassezia furfur*, *Microsporum* spp., *Neotestudina rosatii*, *Onychocola canadensis*, *Paracoccidioides brasiliensis*, *Phialophora verrucosa*, *Piedraia hortae*, *Piedra iahortae*, *Pityriasis versicolor*, *Pseudallescheria boydii*, *Pyrenochaeta romeroi*, *Rhizopus arrhizus*, *Scopulariopsis brevicaulis*, *Scytalidium dimidiatum*, *Sporothrix schenckii*, *Trichophyton* spp., *Trichosporon* spp., *Zygomycete fungi*, *Absidia corymbifera*, *Rhizomucor pusillus* and *Rhizopus arrhizus*. Thus, representative fungal antigens that can be used in the compositions and methods of the present invention include, but are not limited to, candida fungal antigen components; histoplasma fungal antigens such as heat shock protein 60 (HSP60) and other histoplasma fungal antigen components; cryptococcal fungal antigens such as capsular polysaccharides and other cryptococcal fungal antigen components; coccidioides fungal antigens such as spherule antigens and other coccidioides fungal antigen components; and tinea fungal antigens such as trichophytin and other coccidioides fungal antigen components.

[0110] Illustrative examples of bacteria include bacteria that are responsible for diseases including, but not restricted to, diphtheria (*e.g.*, *Corynebacterium diphtheria*), pertussis (*e.g.*, *Bordetella pertussis*, GenBank Accession No. M35274), tetanus (*e.g.*, *Clostridium tetani*, GenBank Accession No. M64353), tuberculosis (*e.g.*, *Mycobacterium tuberculosis*), bacterial pneumonias (*e.g.*, *Haemophilus influenzae.*), cholera (*e.g.*, *Vibrio cholerae*), anthrax (*e.g.*, *Bacillus anthracis*), typhoid, plague, shigellosis (*e.g.*, *Shigella dysenteriae*), botulism (*e.g.*, *Clostridium botulinum*), salmonellosis (*e.g.*, GenBank

Accession No. L03833), peptic ulcers (*e.g.*, Helicobacter pylori), Legionnaire's Disease, Lyme disease (*e.g.*, GenBank Accession No. U59487), Other pathogenic bacteria include Escherichia coli, Clostridium perfringens, Pseudomonas aeruginosa, Staphylococcus aureus and Streptococcus pyogenes. Thus, bacterial antigens which can be used in the compositions and methods of the invention include, but are not limited to: pertussis bacterial antigens such as pertussis toxin, filamentous hemagglutinin, pertactin, F M2, FIM3, adenylate cyclase and other pertussis bacterial antigen components; diphtheria bacterial antigens such as diphtheria toxin or toxoid and other diphtheria bacterial antigen components; tetanus bacterial antigens such as tetanus toxin or toxoid and other tetanus bacterial antigen components, streptococcal bacterial antigens such as M proteins and other streptococcal bacterial antigen components; gram-negative bacilli bacterial antigens such as lipopolysaccharides and other gram-negative bacterial antigen components; Mycobacterium tuberculosis bacterial antigens such as mycolic acid, heat shock protein 65 (HSP65), the 30kDa major secreted protein, antigen 85A and other mycobacterial antigen components; Helicobacter pylori bacterial antigen components, pneumococcal bacterial antigens such as pneumolysin, pneumococcal capsular polysaccharides and other pneumococcal bacterial antigen components; *Haemophilus influenza* bacterial antigens such as capsular polysaccharides and other *Haemophilus influenza* bacterial antigen components; anthrax bacterial antigens such as anthrax protective antigen and other anthrax bacterial antigen components; rickettsiae bacterial antigens such as rompA and other rickettsiae bacterial antigen component. Also included with the bacterial antigens described herein are any other bacterial, mycobacterial, mycoplasmal, rickettsial, or chlamydial antigens.

[0111] Illustrative examples of protozoa include protozoa that are responsible for diseases including, but not limited to, malaria (*e.g.*, GenBank Accession No. X53832), hookworm, onchocerciasis (*e.g.*, GenBank Accession No. M27807), schistosomiasis (*e.g.*, GenBank Accession No. LOS 198), toxoplasmosis, trypanosomiasis, leishmaniasis, giardiasis (GenBank Accession No. M33641), amoebiasis, filariasis (*e.g.*, GenBank Accession No. J03266), borreliosis, and trichinosis. Thus, protozoal antigens which can be used in the compositions and methods of the invention include, but are not limited to: plasmodium falciparum antigens such as merozoite surface antigens, sporozoite surface antigens, circumsporozoite antigens, gametocyte/gamete surface antigens, blood-stage antigen pf 155/RESA and other plasmodial antigen components; toxoplasma antigens such as SAG-1, p30 and other toxoplasma antigen components; schistosoma antigens such as glutathione-S-transferase, paramyosin, and other schistosomal antigen components; leishmania major and

other leishmaniae antigens such as gp63, lipophosphoglycan and its associated protein and other leishmanial antigen components; and trypanosoma cruzi antigens such as the 75-77kDa antigen, the 56kDa antigen and other trypanosomal antigen components.

[0112] The present invention also contemplates toxin components as antigens, illustrative examples of which include staphylococcal enterotoxins, toxic shock syndrome toxin; retroviral antigens (*e.g.*, antigens derived from HIV), streptococcal antigens, staphylococcal enterotoxin-A (SEA), staphylococcal enterotoxin-B (SEB), staphylococcal enterotoxin$_{I\text{-}3}$ ($SE_{1\text{-}3}$), staphylococcal enterotoxin-D (SED), staphylococcal enterotoxin-E (SEE) as well as toxins derived from mycoplasma, mycobacterium, and herpes viruses.

6. Construction of synthetic polynucleotides

[0113] Replacement of one codon for another can be achieved using standard methods known in the art. For example codon modification of a parent polynucleotide can be effected using several known mutagenesis techniques including, for example, oligonucleotide-directed mutagenesis, mutagenesis with degenerate oligonucleotides, and region-specific mutagenesis. Exemplary *in vitro* mutagenesis techniques are described for example in U.S. Patent Nos. 4,184,917, 4,321,365 and 4,351,901 or in the relevant sections of Ausubel, *et al.* (CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, Inc. 1997) and of Sambrook, *et al.*, (MOLECULAR CLONING. A LABORATORY MANUAL, Cold Spring Harbor Press, 1989). Instead of *in vitro* mutagenesis, the synthetic polynucleotide can be synthesized *de novo* using readily available machinery as described, for example, in U.S. Patent No 4,293,652. However, it should be noted that the present invention is not dependent on, and not directed to, any one particular technique for constructing the synthetic polynucleotide.

[0114] The parent polynucleotide is suitably a natural gene. However, it is possible that the parent polynucleotide is not naturally-occurring but has been engineered using recombinant techniques. Parent polynucleotides can be obtained from any suitable source, such as from eukaryotic or prokaryotic organisms, including but not limited to mammals or other animals, and pathogenic organisms such as yeasts, bacteria, protozoa and viruses.

[0115] The invention also contemplates synthetic polynucleotides encoding one or more desired portions of a target antigen. In some embodiments, the synthetic polynucleotide encodes at least about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 40, 50, 60, 70, 80, 90, 100, 120, 150, 300, 400, 500, 600, 700, 800, 900 or 1000, or even at

least about 2000, 3000, 4000 or 5000 contiguous amino acid residues, or almost up to the total number of amino acids present in a full-length target antigen. In some embodiments, the synthetic polynucleotide encodes a plurality of portions of the target antigen, wherein the portions are the same or different. In illustrative examples of this type, the synthetic polynucleotide encodes a multi-epitope fusion protein. A number of factors can influence the choice of portion size. For example, the size of individual portions encoded by the synthetic polynucleotide can be chosen such that it includes, or corresponds to the size of, T cell epitopes and/or B cell epitopes, and their processing requirements. Practitioners in the art will recognize that class I-restricted T cell epitopes are typically between 8 and 10 amino acid residues in length and if placed next to unnatural flanking residues, such epitopes can generally require 2 to 3 natural flanking amino acid residues to ensure that they are efficiently processed and presented. Class II-restricted T cell epitopes usually range between 12 and 25 amino acid residues in length and may not require natural flanking residues for efficient proteolytic processing although it is believed that natural flanking residues may play a role. Another important feature of class II-restricted epitopes is that they generally contain a core of 9-10 amino acid residues in the middle which bind specifically to class II MHC molecules with flanking sequences either side of this core stabilizing binding by associating with conserved structures on either side of class II MHC antigens in a sequence independent manner. Thus the functional region of class II-restricted epitopes is typically less than about 15 amino acid residues long. The size of linear B cell epitopes and the factors effecting their processing, like class II-restricted epitopes, are quite variable although such epitopes are frequently smaller in size than 15 amino acid residues. From the foregoing, it is advantageous, but not essential, that the size of individual portions of the target antigen is at least 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30 amino acid residues. Suitably, the size of individual portions is no more than about 500, 200, 100, 80, 60, 50, 40 amino acid residues. In certain advantageous embodiments, the size of individual portions is sufficient for presentation by an antigen-presenting cell of a T cell and/or a B cell epitope contained within the peptide.

[0116] As will be appreciated by those of skill in the art, it is generally not necessary to immunize with a polypeptide that shares exactly the same amino acid sequence with the target antigen to produce an immune response to that antigen. In some embodiments, therefore, the polypeptide encoded by the synthetic polynucleotide is desirably a variant of at least a portion of the target antigen. "Variant" polypeptides include proteins derived from the target antigen by deletion (so-called truncation) or addition of one or more amino acids to the N-terminal and/or C-terminal end of the target antigen; deletion or addition of one or more

amino acids at one or more sites in the target antigen; or substitution of one or more amino acids at one or more sites in the target antigen. Variant polypeptides encompassed by the present invention will have at least 40%, 50%, 60%, 70%, generally at least 75%, 80%, 85%, typically at least about 90% to 95% or more, and more typically at least about 96%, 97%, 98%, 99% or more sequence similarity or identity with the amino acid sequence of the target antigen or portion thereof as determined by sequence alignment programs described elsewhere herein using default parameters. A variant of a target antigen may differ from that antigen generally by as much 1000, 500, 400, 300, 200, 100, 50 or 20 amino acid residues or suitably by as few as 1-15 amino acid residues, as few as 1-10, such as 6-10, as few as 5, as few as 4, 3, 2, or even 1 amino acid residue.

[0117] Variant polypeptides corresponding to at least a portion of a target antigen may contain conservative amino acid substitutions at various locations along their sequence, as compared to the target antigen amino acid sequence. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art, which can be generally sub-classified as follows:

[0118] Acidic: The residue has a negative charge due to loss of H ion at physiological pH and the residue is attracted by aqueous solution so as to seek the surface positions in the conformation of a peptide in which it is contained when the peptide is in aqueous medium at physiological pH. Amino acids having an acidic side chain include glutamic acid and aspartic acid.

[0119] Basic: The residue has a positive charge due to association with H ion at physiological pH or within one or two pH units thereof (*e.g.*, histidine) and the residue is attracted by aqueous solution so as to seek the surface positions in the conformation of a peptide in which it is contained when the peptide is in aqueous medium at physiological pH. Amino acids having a basic side chain include arginine, lysine and histidine.

[0120] Charged: The residues are charged at physiological pH and, therefore, include amino acids having acidic or basic side chains (*i.e.*, glutamic acid, aspartic acid, arginine, lysine and histidine).

[0121] Hydrophobic: The residues are not charged at physiological pH and the residue is repelled by aqueous solution so as to seek the inner positions in the conformation of a peptide in which it is contained when the peptide is in aqueous medium. Amino acids

having a hydrophobic side chain include tyrosine, valine, isoleucine, leucine, methionine, phenylalanine and tryptophan.

[0122] Neutral/polar: The residues are not charged at physiological pH, but the residue is not sufficiently repelled by aqueous solutions so that it would seek inner positions in the conformation of a peptide in which it is contained when the peptide is in aqueous medium. Amino acids having a neutral/polar side chain include asparagine, glutamine, cysteine, histidine, serine and threonine.

[0123] This description also characterizes certain amino acids as "small" since their side chains are not sufficiently large, even if polar groups are lacking, to confer hydrophobicity. With the exception of proline, "small" amino acids are those with four carbons or less when at least one polar group is on the side chain and three carbons or less when not. Amino acids having a small side chain include glycine, serine, alanine and threonine. The gene-encoded secondary amino acid proline is a special case due to its known effects on the secondary conformation of peptide chains. The structure of proline differs from all the other naturally-occurring amino acids in that its side chain is bonded to the nitrogen of the α-amino group, as well as the α-carbon. Several amino acid similarity matrices (*e.g.*, PAM120 matrix and PAM250 matrix as disclosed for example by Dayhoff *et al.* (1978) A model of evolutionary change in proteins. Matrices for determining distance relationships *In* M. O. Dayhoff, (ed.), Atlas of protein sequence and structure, Vol. 5, pp. 345-358, National Biomedical Research Foundation, Washington DC; and by Gonnet *et al.*, 1992, *Science* **256**(5062): 144301445), however, include proline in the same group as glycine, serine, alanine and threonine. Accordingly, for the purposes of the present invention, proline is classified as a "small" amino acid.

[0124] The degree of attraction or repulsion required for classification as polar or nonpolar is arbitrary and, therefore, amino acids specifically contemplated by the invention have been classified as one or the other. Most amino acids not specifically named can be classified on the basis of known behavior.

[0125] Amino acid residues can be further sub-classified as cyclic or noncyclic, and aromatic or nonaromatic, self-explanatory classifications with respect to the side-chain substituent groups of the residues, and as small or large. The residue is considered small if it contains a total of four carbon atoms or less, inclusive of the carboxyl carbon, provided an additional polar substituent is present; three or less if not. Small residues are, of course, always nonaromatic. Dependent on their structural properties, amino acid residues may fall in

two or more classes. For the naturally-occurring protein amino acids, sub-classification according to the this scheme is presented in the Table 10.

TABLE 10

Original Residue	Exemplary Substitutions
Ala	Ser
Arg	Lys
Asn	Gln, His
Asp	Glu
Cys	Ser
Gln	Asn
Glu	Asp
Gly	Pro
His	Asn, Gln
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile,
Phe	Met, Leu, Tyr
Ser	Thr
Thr	Ser
Trp	Tyr
Tyr	Trp, Phe
Val	Ile, Leu

[0126] Conservative amino acid substitution also includes groupings based on side chains. For example, a group of amino acids having aliphatic side chains is glycine, alanine, valine, leucine, and isoleucine; a group of amino acids having aliphatic-hydroxyl side chains is serine and threonine; a group of amino acids having amide-containing side chains is asparagine and glutamine; a group of amino acids having aromatic side chains is phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains is

lysine, arginine, and histidine; and a group of amino acids having sulfur-containing side chains is cysteine and methionine. For example, it is reasonable to expect that replacement of a leucine with an isoleucine or valine, an aspartate with a glutamate, a threonine with a serine, or a similar replacement of an amino acid with a structurally related amino acid will not have a major effect on the properties of the resulting variant polypeptide. Conservative substitutions are shown in Table 11 below under the heading of exemplary substitutions. More preferred substitutions are shown under the heading of preferred substitutions. Amino acid substitutions falling within the scope of the invention, are, in general, accomplished by selecting substitutions that do not differ significantly in their effect on maintaining (a) the structure of the peptide backbone in the area of the substitution, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. After the substitutions are introduced, the variants are screened for biological activity.

TABLE 11

EXEMPLARY AND PREFERRED AMINO ACID SUBSTITUTIONS

Original Residue	Exemplary Substitutions	Preferred Substitutions
Ala	Val, Leu, Ile	Val
Arg	Lys, Gln, Asn	Lys
Asn	Gln, His, Lys, Arg	Gln
Asp	Glu	Glu
Cys	Ser	Ser
Gln	Asn, His, Lys,	Asn
Glu	Asp, Lys	Asp
Gly	Pro	Pro
His	Asn, Gln, Lys, Arg	Arg
Ile	Leu, Val, Met, Ala, Phe, Norleu	Leu
Leu	Norleu, Ile, Val, Met, Ala, Phe	Ile
Lys	Arg, Gln, Asn	Arg
Met	Leu, Ile, Phe	Leu
Phe	Leu, Val, Ile, Ala	Leu

Original Residue	Exemplary Substitutions	Preferred Substitutions
Pro	Gly	Gly
Ser	Thr	Thr
Thr	Ser	Ser
Trp	Tyr	Tyr
Tyr	Trp, Phe, Thr, Ser	Phe
Val	Ile, Leu, Met, Phe, Ala, Norleu	Leu

[0127] Alternatively, similar amino acids for making conservative substitutions can be grouped into three categories based on the identity of the side chains. The first group includes glutamic acid, aspartic acid, arginine, lysine, histidine, which all have charged side chains; the second group includes glycine, serine, threonine, cysteine, tyrosine, glutamine, asparagine; and the third group includes leucine, isoleucine, valine, alanine, proline, phenylalanine, tryptophan, methionine, as described in Zubay, G., *Biochemistry*, third edition, Wm.C. Brown Publishers (1993).

[0128] The invention further contemplates a chimeric construct comprising a synthetic polynucleotide of the invention, which is operably linked to a regulatory polynucleotide. The regulatory polynucleotide suitably comprises transcriptional and/or translational control sequences, which will be compatible for expression in the organism of interest or in cells of that organism. Typically, the transcriptional and translational regulatory control sequences include, but are not limited to, a promoter sequence, a 5' non-coding region, a *cis*-regulatory region such as a functional binding site for transcriptional regulatory protein or translational regulatory protein, an upstream open reading frame, ribosomal-binding sequences, transcriptional start site, translational start site, and/or nucleotide sequence which encodes a leader sequence, termination codon, translational stop site and a 3' non-translated region. Constitutive or inducible promoters as known in the art are contemplated by the invention. The promoters may be either naturally occurring promoters, or hybrid promoters that combine elements of more than one promoter. Promoter sequences contemplated by the present invention may be native to the organism of interest or may be derived from an alternative source, where the region is functional in the chosen organism. The choice of promoter will differ depending on the intended host or cell or tissue type. For example, promoters which could be used for expression in mammals include the

metallothionein promoter, which can be induced in response to heavy metals such as cadmium, the β-actin promoter as well as viral promoters such as the SV40 large T antigen promoter, human cytomegalovirus (CMV) immediate early (IE) promoter, Rous sarcoma virus LTR promoter, the mouse mammary tumor virus LTR promoter, the adenovirus major late promoter (Ad MLP), the herpes simplex virus promoter, and a HPV promoter, particularly the HPV upstream regulatory region (URR), among others. All these promoters are well described and readily available in the art.

[0129] Enhancer elements may also be used herein to increase expression levels of the mammalian constructs. Examples include the SV40 early gene enhancer, as described for example in Dijkema *et al.* (1985, EMBO J. 4:761), the enhancer/promoter derived from the long terminal repeat (LTR) of the Rous Sarcoma Virus, as described for example in Gorman *et al.*, (1982, Proc. Natl. Acad. Sci. USA 79:6777) and elements derived from human CMV, as described for example in Boshart *et al.* (1985, Cell 41:521), such as elements included in the CMV intron A sequence.

[0130] The chimeric construct may also comprise a 3' non-translated sequence. A 3' non-translated sequence refers to that portion of a gene comprising a DNA segment that contains a polyadenylation signal and any other regulatory signals capable of effecting mRNA processing or gene expression. The polyadenylation signal is characterized by effecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor. Polyadenylation signals are commonly recognized by the presence of homology to the canonical form 5' AATAAA-3' although variations are not uncommon. The 3' non-translated regulatory DNA sequence preferably includes from about 50 to 1,000 nts and may contain transcriptional and translational termination sequences in addition to a polyadenylation signal and any other regulatory signals capable of effecting mRNA processing or gene expression.

[0131] In some embodiments, the chimeric construct further contains a selectable marker gene to permit selection of cells containing the construct. Selection genes are well known in the art and will be compatible for expression in the cell of interest.

[0132] It will be understood, however, that expression of protein-encoding polynucleotides in heterologous systems is now well known, and the present invention is not directed to or dependent on any particular vector, transcriptional control sequence or technique for expression of the polynucleotides. Rather, synthetic polynucleotides prepared according to the methods set forth herein may be introduced into a mammal in any suitable

manner in the form of any suitable construct or vector, and the synthetic polynucleotides may be expressed with known transcription regulatory elements in any conventional manner.

[0133] In addition, chimeric constructs can be constructed that include sequences coding for adjuvants. Particularly suitable are detoxified mutants of bacterial ADP-ribosylating toxins, for example, diphtheria toxin, pertussis toxin (PT), cholera toxin (CT), *Escherichia coli* heat-labile toxins (LT1 and LT2), *Pseudomonas* endotoxin A, *Clostridium botulinum* C2 and C3 toxins, as well as toxins from *C. perfringens*, *C. spiriforma* and *C. difficile*. In some embodiments, the chimeric constructs include coding sequences for detoxified mutants of *E. coli* heat-labile toxins, such as the LT-K63 and LT-R72 detoxified mutants, described in U.S. Pat. No. 6,818,222. In some embodiments, the adjuvant is a protein-destabilising element, which increases processing and presentation of the polypeptide that corresponds to at least a portion of the target antigen through the class I MHC pathway, thereby leading to enhanced cell-mediated immunity against the polypeptide. Illustrative protein-destabilising elements include intracellular protein degradation signals or degrons which may be selected without limitation from a destabilising amino acid at the amino-terminus of a polypeptide of interest, a PEST region or a ubiquitin. For example, the coding sequence for the polypeptide can be modified to include a destabilising amino acid at its amino-terminus so that the protein so modified is subject to the N-end rule pathway as disclosed, for example, by Bachmair *et al.* in U.S. Patent Serial No. 5,093,242 and by Varshavsky *et al.* in U.S. Patent Serial No. 5,122,463. In some embodiments, the destabilising amino acid is selected from isoleucine and glutamic acid, especially from histidine tyrosine and glutamine, and more especially from aspartic acid, asparagine, phenylalanine, leucine, tryptophan and lysine. In certain embodiments, the destabilising amino acid is arginine. In some proteins, the amino-terminal end is obscured as a result of the protein's conformation (*i.e.*, its tertiary or quaternary structure). In these cases, more extensive alteration of the amino-terminus may be necessary to make the protein subject to the N-end rule pathway. For example, where simple addition or replacement of the single amino-terminal residue is insufficient because of an inaccessible amino-terminus, several amino acids (including lysine, the site of ubiquitin joining to substrate proteins) may be added to the original amino-terminus to increase the accessibility and/or segmental mobility of the engineered amino terminus. In some embodiments, a nucleic acid sequence encoding the amino-terminal region of the polypeptide can be modified to introduce a lysine residue in an appropriate context. This can be achieved most conveniently by employing DNA constructs encoding "universal destabilising segments". A universal destabilising segment comprises a nucleic acid construct

which encodes a polypeptide structure, preferably segmentally mobile, containing one or more lysine residues, the codons for lysine residues being positioned within the construct such that when the construct is inserted into the coding sequence of the protein-encoding synthetic polynucleotide, the lysine residues are sufficiently spatially proximate to the amino-terminus of the encoded protein to serve as the second determinant of the complete amino-terminal degradation signal. The insertion of such constructs into the 5' portion of a polypeptide-encoding synthetic polynucleotide would provide the encoded polypeptide with a lysine residue (or residues) in an appropriate context for destabilization. In other embodiments, the polypeptide is modified to contain a PEST region, which is rich in an amino acid selected from proline, glutamic acid, serine and threonine, which region is optionally flanked by amino acids comprising electropositive side chains. In this regard, it is known that amino acid sequences of proteins with intracellular half-lives less than about 2 hours contain one or more regions rich in proline (P), glutamic acid (E), serine (S), and threonine (T) as for example shown by Rogers et al. (1986, Science 234 (4774): 364-368). In still other embodiments, the polypeptide is conjugated to a ubiquitin or a biologically active fragment thereof, to produce a modified polypeptide whose rate of intracellular proteolytic degradation is increased, enhanced or otherwise elevated relative to the unmodified polypeptide.

[0134] One or more adjuvant polypeptides may be co-expressed with an 'antigenic' polypeptide that corresponds to at least a portion of the target antigen. In certain embodiments, adjuvant and antigenic polypeptides may be co-expressed in the form of a fusion protein comprising one or more adjuvant polypeptides and one or more antigenic polypeptides. Alternatively, adjuvant and antigenic polypeptides may be co-expressed as separate proteins.

[0135] Furthermore, chimeric constructs can be constructed that include chimeric antigen-coding gene sequences, encoding, *e.g.*, multiple antigens/epitopes of interest, for example derived from a single or from more than one target antigen. In certain embodiments, multi-cistronic cassettes (*e.g.*, bi-cistronic cassettes) can be constructed allowing expression of multiple adjuvants and/or antigenic polypeptides from a single mRNA using, for example, the EMCV IRES, or the like. In other embodiments, adjuvants and/or antigenic polypeptides can be encoded on separate coding sequences that are operably connected to independent transcription regulatory elements.

[0136] In some embodiments, the chimeric constructs of the invention are in the form of expression vectors which are suitably selected from self-replicating extra-

chromosomal vectors (*e.g.*, plasmids) and vectors that integrate into a host genome. In illustrative examples of this type, the expression vectors are viral vectors, such as simian virus 40 (SV40) or bovine papilloma virus (BPV), which has the ability to replicate as extra-chromosomal elements (Eukaryotic Viral Vectors, Cold Spring Harbor Laboratory, Gluzman ed., 1982; Sarver *et al.*, 1981, Mol. Cell. Biol. 1:486). Viral vectors include retroviral (lentivirus), adeno-associated virus (see, *e.g.*, Okada, 1996, Gene Ther. 3:957-964; Muzyczka, 1994, J. Clin. Invst. 94:1351; U.S. Pat. Nos. 6,156,303; 6,143,548 5,952,221, describing AAV vectors; see also U.S. Pat. Nos. 6,004,799; 5,833,993), adenovirus (see, *e.g.*, U.S. Pat. Nos. 6,140,087; 6,136,594; 6,133,028; 6,120,764), reovirus, herpesvirus, rotavirus genomes etc., modified for introducing and directing expression of a polynucleotide or transgene in cells. Retroviral vectors can include those based upon murine leukemia virus (see, *e.g.*, U.S. Pat. No. 6,132,731), gibbon ape leukemia virus (see, *e.g.*, U.S. Pat. No. 6,033,905), simian immuno-deficiency virus, human immuno-deficiency virus (see, *e.g.*, U.S. Pat. No. 5,985,641), and combinations thereof.

[0137] Vectors also include those that efficiently deliver genes to animal cells *in vivo* (*e.g.*, stem cells) (see, *e.g.*, U.S. Pat. Nos. 5,821,235 and 5,786,340; Croyle *et al.*, 1998, Gene Ther. 5:645; Croyle *et al.*, *1998*, Pharm. Res. *15:1348*; Croyle *et al.*, *1998*, Hum. Gene Ther. 9:561; Foreman *et al.*, 1998, Hum. Gene Ther. 9:1313 ; Wirtz *et al.*, 1999, Gut 44:800). Adenoviral and adeno-associated viral vectors suitable for *in vivo* delivery are described, for example, in U.S. Pat. Nos. 5,700,470, 5,731,172 and 5,604,090. Additional vectors suitable for *in vivo* delivery include herpes simplex virus vectors (see, *e.g.*, U.S. Pat. No. 5,501,979), retroviral vectors (see, *e.g.*, U.S. Pat. Nos. 5,624,820, 5,693,508 and 5,674,703; and WO92/05266 and WO92/14829), bovine papilloma virus (BPV) vectors (see, *e.g.*, U.S. Pat. No. 5,719,054), CMV-based vectors (see, *e.g.*, U.S. Pat. No. 5,561,063) and parvovirus, rotavirus and Norwalk virus vectors. Lentiviral vectors are useful for infecting dividing as well as non-dividing cells (see, *e.g.*, U.S. Pat. No. 6,013,516).

[0138] Additional viral vectors which will find use for delivering the nucleic acid molecules encoding the antigens of interest include those derived from the pox family of viruses, including vaccinia virus and avian poxvirus. By way of example, vaccinia virus recombinants expressing the chimeric constructs can be constructed as follows. The antigen coding sequence is first inserted into an appropriate vector so that it is adjacent to a vaccinia promoter and flanking vaccinia DNA sequences, such as the sequence encoding thymidine kinase (TK). This vector is then used to transfect cells that are simultaneously infected with

vaccinia. Homologous recombination serves to insert the vaccinia promoter plus the gene encoding the coding sequences of interest into the viral genome. The resulting TK- recombinant can be selected by culturing the cells in the presence of 5-bromodeoxyuridine and picking viral plaques resistant thereto.

[0139] Alternatively, avipoxviruses, such as the fowlpox and canarypox viruses, can also be used to deliver the genes. Recombinant avipox viruses, expressing immunogens from mammalian pathogens, are known to confer protective immunity when administered to non-avian species. The use of an avipox vector is particularly desirable in human and other mammalian species since members of the avipox genus can only productively replicate in susceptible avian species and therefore are not infective in mammalian cells. Methods for producing recombinant avipoxviruses are known in the art and employ genetic recombination, as described above with. respect to the production of vaccinia viruses. See, *e.g.*, WO 91/12882; WO 89/03429; and WO 92/03545.

[0140] Molecular conjugate vectors, such as the adenovirus chimeric vectors described in Michael et al., J. Biol. Chem. (1993) 268:6866-6869 and Wagner et al., Proc. Natl. Acad. Sci. USA (1992) 89:6099-6103, can also be used for gene delivery.

[0141] Members of the Alphavirus genus, such as, but not limited to, vectors derived from the Sindbis virus (SIN), Semliki Forest virus (SFV), and Venezuelan Equine Encephalitis virus (VEE), will also find use as viral vectors for delivering the chimeric constructs of the present invention. For a description of Sindbis-virus derived vectors useful for the practice of the instant methods, see, Dubensky *et al.* (1996, J. Virol. 70:508-519; and International Publication Nos. WO 95/07995, WO 96/17072); as well as, Dubensky, Jr., T. W., *et al.*, U.S. Pat. No. 5,843,723, and Dubensky, Jr., T. W., U.S. Pat. No. 5,789,245. Exemplary vectors of this type are chimeric alphavirus vectors comprised of sequences derived from Sindbis virus and Venezuelan equine encephalitis virus. See, *e.g.*, Perri *et al.* (2003, J. Virol. 77: 10394-10403) and International Publication Nos. WO 02/099035, WO 02/080982, WO 01/81609, and WO 00/61772.

[0142] In other illustrative embodiments, lentiviral vectors are employed to deliver a chimeric construct of the inevntion into selected cells or tissues. Typically, these vectors comprise a 5' lentiviral LTR, a tRNA binding site, a packaging signal, a promoter operably linked to one or more genes of interest, an origin of second strand DNA synthesis and a 3' lentiviral LTR, wherein the lentiviral vector contains a nuclear transport element. The nuclear transport element may be located either upstream (5') or downstream (3') of a coding

sequence of interest (for example, a synthetic Gag or Env expression cassette of the present invention). A wide variety of lentiviruses may be utilized within the context of the present invention, including for example, lentiviruses selected from the group consisting of HIV, HIV-1, HIV-2, FIV, BIV, EIAV, MVV, CAEV, and SIV. Illustrative examples of lentiviral vectors are described in PCT Publication Nos. WO 00/66759, WO 00/00600, WO 99/24465, WO 98/51810, WO 99/51754, WO 99/31251, WO 99/30742, and WO 99/15641. Desirably, a third generation SIN lentivirus is used. Commercial suppliers of third generation SIN (self-inactivating) lentiviruses include Invitrogen (ViraPower Lentiviral Expression System). Detailed methods for construction, transfection, harvesting, and use of lentiviral vectors are given, for example, in the Invitrogen technical manual "ViraPower Lentiviral Expression System version B 050102 25-0501", available at http://www.invitrogen.com/Content/Tech-Online/molecular_biology/manuals_p-ps/virapower_lentiviral_system_man.pdf. Lentiviral vectors have emerged as an efficient method for gene transfer. Improvements in biosafety characteristics have made these vectors suitable for use at biosafety level 2 (BL2). A number of safety features are incorporated into third generation SIN (self-inactivating) vectors. Deletion of the viral 3' LTR U3 region results in a provirus that is unable to transcribe a full length viral RNA. In addition, a number of essential genes are provided in trans, yielding a viral stock that is capable of but a single round of infection and integration. Lentiviral vectors have several advantages, including: 1) pseudotyping of the vector using amphotropic envelope proteins allows them to infect virtually any cell type; 2) gene delivery to quiescent, post mitotic, differentiated cells, including neurons, has been demonstrated; 3) their low cellular toxicity is unique among transgene delivery systems; 4) viral integration into the genome permits long term transgene expression; 5) their packaging capacity (6-14 kb) is much larger than other retroviral, or adeno-associated viral vectors. In a recent demonstration of the capabilities of this system, lentiviral vectors expressing GFP were used to infect murine stem cells resulting in live progeny, germline transmission, and promoter-, and tissue-specific expression of the reporter (Ailles, L. E. and Naldini, L., HIV-1-Derived Lentiviral Vectors. In: Trono,D. (Ed.), Lentiviral Vectors, Springer-Verlag, Berlin, Heidelberg, New York, 2002, pp. 31-52). An example of the current generation vectors is outlined in FIG. 2 of a review by Lois *et al.* (2002, Science, 295 868-872).

[0143] The chimeric construct can also be delivered without a vector. For example, the chimeric construct can be packaged as DNA or RNA in liposomes prior to delivery to the subject or to cells derived therefrom. Lipid encapsulation is generally accomplished using liposomes which are able to stably bind or entrap and retain nucleic acid. The ratio of

condensed DNA to lipid preparation can vary but will generally be around 1:1 (mg DNA:micromoles lipid), or more of lipid. For a review of the use of liposomes as carriers for delivery of nucleic acids, see, Hug and Sleight, (1991, Biochim. Biophys. Acta. 1097:1-17); and Straubinger *et al.*, in Methods of Enzymology (1983), Vol. 101, pp. 512-527.

[0144] Liposomal preparations for use in the present invention include cationic (positively charged), anionic (negatively charged) and neutral preparations, with cationic liposomes particularly preferred. Cationic liposomes have been shown to mediate intracellular delivery of plasmid DNA (Felgner *et al.*, 1987, Proc. Natl. Acad. Sci. USA 84:7413-7416); mRNA (Malone *et al.*, 1989, Proc. Natl. Acad. Sci. USA 86:6077-6081); and purified transcription factors (Debs *et al.*, 1990, J. Biol. Chem. 265:10189-10192), in functional form.

[0145] Cationic liposomes are readily available. For example, N[1-2,3-dioleyloxy)propyl]-N,N,N-triethylammonium (DOTMA) liposomes are available under the trademark Lipofectin, from GIBCO BRL, Grand Island, N.Y. (See, also, Felgner *et al.*, 1987, Proc. Natl. Acad. Sci. USA 84:7413-7416). Other commercially available lipids include (DDAB/DOPE) and DOTAP/DOPE (Boerhinger). Alternative cationic liposomes can be prepared from readily available materials using techniques well known in the art. See, *e.g.*, Szoka *et al.*, 1978, Proc. Natl. Acad. Sci. USA 75:4194-4198; PCT Publication No. WO 90/11092 for a description of the synthesis of DOTAP (1,2-bis(oleoyloxy)-3-(trimethylammonio)propane) liposomes.

[0146] Similarly, anionic and neutral liposomes are readily available, such as, from Avanti Polar Lipids (Birmingham, Ala.), or can be easily prepared using readily available materials. Such materials include phosphatidyl choline, cholesterol, phosphatidyl ethanolamine, dioleoylphosphatidyl choline (DOPC), dioleoylphosphatidyl glycerol (DOPG), dioleoylphoshatidyl ethanolamine (DOPE), among others. These materials can also be mixed with the DOTMA and DOTAP starting materials in appropriate ratios. Methods for making liposomes using these materials are well known in the art.

[0147] The liposomes can comprise multilammelar vesicles (MLVs), small unilamellar vesicles (SUVs), or large unilamellar vesicles (LUVs). The various liposome-nucleic acid complexes are prepared using methods known in the art. See, *e.g.*, Straubinger *et al.*, in METHODS OF IMMUNOLOGY (1983), Vol. 101, pp. 512-527; Szoka *et al.*, 1978, Proc. Natl. Acad. Sci. USA 75:4194-4198; Papahadjopoulos *et al.*, 1975, Biochim. Biophys. Acta 394:483; Wilson *et al.*, 1979, Cell 17:77); Deamer and Bangham, 1976, Biochim. Biophys. Acta 443:629; Ostro *et al.*, 1977, Biochem. Biophys. Res. Commun. 76:836; Fraley

et al., 1979, Proc. Natl. Acad. Sci. USA 76:3348); Enoch and Strittmatter, 1979, Proc. Natl. Acad. Sci. USA 76:145); Fraley *et al.*, 1980, J. Biol. Chem. 255:10431; Szoka and Papahadjopoulos, 1978, Proc. Natl. Acad. Sci. USA 75:145; and Schaefer-Ridder *et al.*, 1982, Science 215:166.

[0148] The chimeric construct can also be delivered in cochleate lipid compositions similar to those described by Papahadjopoulos *et al.*, 1975, Biochem. Biophys. Acta. 394:483-491. See, also, U.S. Pat. Nos. 4,663,161 and 4,871,488.

[0149] The chimeric construct may also be encapsulated, adsorbed to, or associated with, particulate carriers. Such carriers present multiple copies-of a selected chimeric construct to the immune system. The particles can be taken up by professional antigen presenting cells such as macrophages and dendritic cells, and/or can enhance antigen presentation through other mechanisms such as stimulation of cytokine release. Examples of particulate carriers include those derived from polymethyl methacrylate polymers, as well as microparticles derived from poly(lactides) and poly(lactide-co-glycolides), known as PLG. See, *e.g.*, Jeffery *et al.*, 1993, Pharm. Res. 10:362-368; McGee J. P., *et al.*, 1997, J Microencapsul. 14(2):197-210; O'Hagan D. T., *et al.*, 1993, Vaccine 11(2):149-54.

[0150] Furthermore, other particulate systems and polymers can be used for the *in vivo* delivery of the chimeric construct. For example, polymers such as polylysine, polyarginine, polyornithine, spermine, spermidine, as well as conjugates of these molecules, are useful for transferring a nucleic acid of interest. Similarly, DEAE dextran-mediated transfection, calcium phosphate precipitation or precipitation using other insoluble inorganic salts, such as strontium phosphate, aluminum silicates including bentonite and kaolin, chromic oxide, magnesium silicate, talc, and the like, will find use with the present methods. See, *e.g.*, Felgner, P. L., Advanced Drug Delivery Reviews (1990) 5:163-187, for a review of delivery systems useful for gene transfer. Peptoids (Zuckerman, R. N., et al., U.S. Pat. No. 5,831,005, issued Nov. 3, 1998) may also be used for delivery of a construct of the present invention.

[0151] Additionally, biolistic delivery systems employing particulate carriers such as gold and tungsten, are especially useful for delivering chimeric constructs of the present invention. The particles are coated with the synthetic expression cassette(s) to be delivered and accelerated to high velocity, generally under a reduced atmosphere, using a gun powder discharge from a "gene gun." For a description of such techniques, and apparatuses useful therefor, see, *e.g.*, U.S. Pat. Nos. 4,945,050; 5,036,006; 5,100,792; 5,179,022; 5,371,015; and

5,478,744. In illustrative examples, gas-driven particle acceleration can be achieved with devices such as those manufactured by PowderMed Pharmaceuticals PLC (Oxford, UK) and PowderMed Vaccines Inc. (Madison, Wis.), some examples of which are described in U.S. Pat. Nos. 5,846,796; 6,010,478; 5,865,796; 5,584,807; and EP Patent No. 0500 799. This approach offers a needle-free delivery approach wherein a dry powder formulation of microscopic particles, such as polynucleotide or polypeptide particles, are accelerated to high speed within a helium gas jet generated by a hand held device, propelling the particles into a target tissue of interest. Other devices and methods that may be useful for gas-driven needle-less injection of compositions of the present invention include those provided by Bioject, Inc. (Portland, Oreg.), some examples of which are described in U.S. Pat. Nos. 4,790,824; 5,064,413; 5,312,335; 5,383,851; 5,399,163; 5,520,639 and 5,993,412.

[0152] Alternatively, micro-cannula- and microneedle-based devices (such as those being developed by Becton Dickinson and others) can be used to administer the chimeric constructs of the invention. Illustrative devices of this type are described in EP 1 092 444 A1, and U.S. application Ser. No. 606,909, filed Jun. 29, 2000. Standard steel cannula can also be used for intra-dermal delivery using devices and methods as described in U.S. Ser. No. 417,671, filed Oct. 14, 1999. These methods and devices include the delivery of substances through narrow gauge (about 30 G) "micro-cannula" with limited depth of penetration, as defined by the total length of the cannula or the total length of the cannula that is exposed beyond a depth-limiting feature. It is within the scope of the present invention that targeted delivery of substances including chimeric constructs can be achieved either through a single microcannula or an array of microcannula (or "microneedles"), for example 3-6 microneedles mounted on an injection device that may include or be attached to a reservoir in which the substance to be administered is contained.

7. Compositions

[0153] The invention also provides compositions, particularly immunomodulating compositions, comprising one or more of the chimeric constructs described herein. The immunomodulating compositions may comprise a mixture of chimeric constructs, which in turn may be delivered, for example, using the same or different vectors or vehicles. Antigens may be administered individually or in combination, in *e.g.*, prophylactic (*i.e.*, to prevent infection or disease) or therapeutic (to treat infection or disease) immunomodulating compositions. The immunomodulating compositions may be given more than once (*e.g.*, a "prime" administration followed by one or more "boosts") to achieve the desired effects. The

same composition can be administered in one or more priming and one or more boosting steps. Alternatively, different compositions can be used for priming and boosting.

[0154] The immunomodulating compositions will generally include one or more "pharmaceutically acceptable excipients or vehicles" such as water, saline, glycerol, ethanol, etc. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles.

[0155] Immunomodulating compositions will typically, in addition to the components mentioned above, comprise one or more "pharmaceutically acceptable carriers." These include any carrier which does not itself induce the production of antibodies harmful to the individual receiving the composition. Suitable carriers typically are large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, and lipid aggregates (such as oil droplets or liposomes). Such carriers are well known to those of ordinary skill in the art. A composition may also contain a diluent, such as water, saline, glycerol, etc. Additionally, an auxiliary substance, such as a wetting or emulsifying agent, pH buffering substance, and the like, may be present. A thorough discussion of pharmaceutically acceptable components is available in Gennaro (2000) Remington: The Science and Practice of Pharmacy. 20th ed., ISBN: 0683306472.

[0156] Pharmaceutically compatible salts can also be used in compositions of the invention, for example, mineral salts such as hydrochlorides, hydrobromides, phosphates, or sulfates, as well as salts of organic acids such as acetates, proprionate, malonates, or benzoates. Especially useful protein substrates are serum albumins, keyhole limpet hemocyanin, immunoglobulin molecules, thyroglobulin, ovalbumin, tetanus toxoid, and other proteins well known to those of skill in the art.

[0157] The chimeric constructs of the invention can also be adsorbed to, entrapped within or otherwise associated with liposomes and particulate carriers such as PLG.

[0158] The chimeric constructs of the present invention are formulated into compositions for delivery to a mammal. These compositions may either be prophylactic (to prevent infection) or therapeutic (to treat disease after infection). The compositions will comprise a "therapeutically effective amount" of the gene of interest such that an amount of the antigen can be produced *in vivo* so that an immune response is generated in the individual to which it is administered. The exact amount necessary will vary depending on the subject

being treated; the age and general condition of the subject to be treated; the capacity of the subject's immune system to synthesize antibodies; the degree of protection desired; the severity of the condition being treated; the particular antigen selected and its mode of administration, among other factors. An appropriate effective amount can be readily determined by one of skill in the art. Thus, a "therapeutically effective amount" will fall in a relatively broad range that can be determined through routine trials.

[0159] Once formulated, the compositions of the invention can be administered directly to the subject (*e.g.*, as described above). Direct delivery of chimeric construct-containing compositions *in vivo* will generally be accomplished with or without vectors, as described above, by injection using either a conventional syringe, needless devices such as Bioject™ or a gene gun, such as the Accell™ gene delivery system (PowderMed Ltd, Oxford, England) or microneedle device. The constructs can be delivered (*e.g.*, injected) either subcutaneously, epidermally, intradermally, intramuscularly, intravenous, intramucosally (such as nasally, rectally and vaginally), intraperitoneally or orally. Delivery of nucleic acid into cells of the epidermis is particularly preferred as this mode of administration provides access to skin-associated lymphoid cells and provides for a transient presence of nucleic acid (*e.g.*, DNA) in the recipient. Other modes of administration include oral ingestion and pulmonary administration, suppositories, needle-less injection, transcutaneous, topical, and transdermal applications. Dosage treatment may be a single dose schedule or a multiple dose schedule.

[0160] In order that the invention may be readily understood and put into practical effect, particular preferred embodiments will now be described by way of the following non-limiting examples.

EXAMPLES

EXAMPLE 1

SYNTHETIC CONSTRUCT SYSTEM FOR DETERMINING THE IMMUNE RESPONSE PREFERENCE OF CODONS IN MAMMALS

Materials and Methods

Primer design/ synthesis and sequence manipulation

[0161] Oligonucleotides for site-directed mutagenesis were designed according to the guidelines included in the mutagenesis kit manuals (Quikchange II Site-directed Mutagenesis kit or Quikchange Multi Site-directed Mutagenesis Kit; Stratagene, La Jolla CA). These primers were synthesized and PAGE purified by Sigma (formerly Proligo).

[0162] Oligonucleotides for whole gene synthesis were designed by eye and synthesized by Sigma (formerly Proligo). The primers were supplied as standard desalted oligos. No additional purification of the oligonucleotides was carried out.

[0163] Sequence manipulation and analysis was carried out using the suite of programs on Biomanager (ANGIS) and various other web-based programs including BLAST at NCBI (http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi), NEBcutter V2.0 from New England Biolabs (http://tools.neb.com/NEBcutter2/index.php), the Translate Tool on ExPASy (http://au.expasy.org/tools/dna.html), and the SignalP 3.0 server (http://www.cbs.dtu.dk/services/SignalP/).

Standard cloning techniques

[0164] Restriction enzyme digests, alkaline phosphatase treatments and ligations were carried out according to the enzyme manufacturers' instructions (various manufacturers including New England Biolabs, Roche and Fermentas).

[0165] Purification of DNA from agarose gels and preparation of mini-prep DNA were carried out using commercial kits (Qiagen, Bio-Rad, Macherey-Nagel).

[0166] Agarose gel electrophoresis, phenol/chloroform extraction of contaminant protein from DNA, ethanol precipitation of DNA and other basic molecular biological procedures were carried out using standard protocols, similar to those described in Current Protocols in Molecular Biology (Ebook available via Wiley InterScience; edited by Ausubel *et al.*).

[0167] Sequencing was carried out by the Australian Genome Research Facility (AGRF, Brisbane).

Whole gene synthesis

[0168] Overlapping ~35-50mer oligonucleotides (Sigma-Proligo) were used to synthesize longer DNA sequences. Restriction enzyme sites were incorporated to facilitate cloning. The method used to synthesize the fragments is based on that given in Smith *et al.* (2003). First, oligonucleotides for the top or bottom strand were mixed and then phosphorylated using T4 polynucleotide kinase (PNK; New England Biolabs). The oligonucleotide mixes were then purified from the PNK by a standard phenol/ chloroform extraction and sodium acetate/ ethanol (NaAc/EtOH) precipitation. Equal volumes of oligonucleotide mixes for the top and bottom strands were then mixed and the oligonucleotides denatured by heating at 95°C for 2 mins. The oligonucleotides were annealed by slowly cooling the sample to 55°C and the annealed oligonucleotides ligated using Taq ligase (New England Biolabs). The resulting fragment was purified by phenol/$CHCl_3$ extraction and NaAc/EtOH precipitation.

[0169] The ends of the fragments were filled in and the fragments then amplified, using the outermost forward and reverse primers, with the Clontech Advantage HF 2 PCR kit (Clontech) according to the manufacturer's instructions. To fill in the ends the following PCR was used: 35 cycles of a denaturation step of 94°C for 15s, a slow annealing step where the temperature was ramped down to 55°C over 7 minutes and then kept at 55°C for 2 min, and an elongation step of 72°C for 6 minutes. A final elongation step for 7min at 72°C was then carried out. The second PCR to amplify the fragment involved: an initial denaturation step at 94°C for 30s, followed by 25 cycles of 94°C for 15s, 55°C 30s and 68°C for 1 min, and a final elongation step of 68°C for 3 mins.

[0170] The fragments were then purified by gel electrophoresis, digested and ligated into the relevant vector. Following transformation of *E. coli* with the ligation mixture, mini-preps were made for multiple colonies and the inserts sequenced. Sometimes it was not possible to isolate clones with entirely correct sequence. In those cases the errors were fixed by single or multi site-directed mutagenesis.

Site-directed mutagenesis

[0171] Mutagenesis was carried out using the Quikchange II Site-directed Mutagenesis kit or Quikchange Multi Site-directed Mutagenesis Kit (Stratagene, La Jolla

CA), with appropriate PAGE (polyacrylamide gel electrophoresis)-purified primers (Sigma), according to the manufacturer's instructions.

Preparation of constructs

[0172] The details of the constructs used to generate the codon preference table are summarized in TABLE 12. All constructs were made using pCDNA3 from Invitrogen and were verified by sequencing prior to use.

TABLE 12

SUMMARY OF SECRETORY E7 CONSTRUCT SERIES 1 AND 2

Construct	AA & Codon	CU of Sec Seq	CU of E7	E7 Protein
Control Constructs				
IgkC1	N/A	wt	wt	non-onc
IgkC2	N/A	mc	mc	non-onc
IgkC3	N/A	wt	wt	onc
IgkC4	N/A	mc	mc	onc
Secretory E7 construct series 1				
IgkS1-1	Ala GCG	wt	wt with all Ala gcg	non-onc
IgkS1-2	Ala GCA	wt	wt with all Ala gca	non-onc
IgkS1-3	Ala GCT	wt	wt with all Ala gct	non-onc
IgkS1-4	Ala GCC	wt	wt with all Ala gcc	non-onc
IgkS1-5	Arg AGG	wt	wt with all Arg agg	non-onc
IgkS1-6	Arg AGA	wt	wt with all Arg aga	non-onc
IgkS1-7	Arg CGG	wt	wt with all Arg cgg	non-onc
IgkS1-8	Arg CGA	wt	wt with all Arg cga	non-onc
IgkS1-9	Arg CGT	wt	wt with all Arg	non-onc

Construct	AA & Codon	CU of Sec Seq	CU of E7	E7 Protein
IgkS1-10	Arg CGC	wt	cgt wt with all Arg cgc	non-onc
IgkS1-11	Asn AAT	wt	wt with all Asn aat	non-onc
IgkS1-12	Asn AAC	wt	wt with all Asn aac	non-onc
IgkS1-13	Asp GAT	wt with all Asp gat	wt with all Asp gat	non-onc
IgkS1-14	Asp GAC	wt with all Asp gac	wt with all Asp gac	non-onc
IgkS1-15	Cys TGT	wt	wt with all Cys tgt	non-onc
IgkS1-16	Cys TGC	wt	wt with all Cys tgc	non-onc
IgkS1-17	Glu GAG	wt with all Glu gag	wt with all Glu gag	non-onc
IgkS1-18	Glu GAA	wt with all Glu gaa	wt with all Glu gaa	non-onc
IgkS1-19	Gln CAG	wt	wt with all Gln cag	non-onc
IgkS1-20	Gln CAA	wt	wt with all Gln caa	non-onc
IgkS1-21	Gly GGG	wt with all Gly ggg	wt with all Gly ggg	non-onc
IgkS1-22	Gly GGA	wt with all Gly gga	wt with all Gly gga	non-onc
IgkS1-23	Gly GGT	wt with all Gly ggt	wt with all Gly ggt	non-onc
IgkS1-24	Gly GGC	wt with all Gly ggc	wt with all Gly ggc	non-onc
IgkS1-25	His CAT	wt	wt with all His cat	non-onc
IgkS1-26	His CAC	wt	wt with all His cac	non-onc

Construct	AA & Codon	CU of Sec Seq	CU of E7	E7 Protein
IgkS1-27	Ile ATA	wt	wt with all Ile ata	non-onc
IgkS1-28	Ile ATT	wt	wt with all Ile att	non-onc
IgkS1-29	Ile ATC	wt	wt with all Ile atc	non-onc
IgkS1-30	Lys AAG	wt	wt with all Lys aag	non-onc
IgkS1-31	Lys AAA	wt	wt with all Lys aaa	non-onc
IgkS1-32	Phe TTT	wt	wt with all Phe ttt	non-onc L15F, L22F
IgkS1-33	Phe TTC	wt	wt with all Phe ttc	non-onc L15F, L22F
IgkS1-34	Ser AGT	wt with all Ser agt	wt with all Ser agt	non-onc
IgkS1-35	Ser AGC	wt with all Ser agc	wt with all Ser agc	non-onc
IgkS1-36	Ser TCG	wt with all Ser tcg	wt with all Ser tcg	non-onc
IgkS1-37	Ser TCA	wt with all Ser tca	wt with all Ser tca	non-onc
IgkS1-38	Ser TCT	wt with all Ser tct	wt with all Ser tct	non-onc
IgkS1-39	Ser TCC	wt	wt with all Ser tcc	non-onc
IgkS1-40	Thr ACG	wt with all Thr acg	wt with all Thr acg	non-onc
IgkS1-41	Thr ACA	wt with all Thr aca	wt with all Thr aca	non-onc
IgkS1-42	Thr ACT	wt with all Thr act	wt with all Thr act	non-onc
IgkS1-43	Thr ACC	wt with all Thr acc	wt with all Thr acc	non-onc
IgkS1-44	Tyr TAT	wt	wt with all Tyr tat	non-onc

Construct	AA & Codon	CU of Sec Seq	CU of E7	E7 Protein
IgkS1-45	Tyr TAC	wt	wt with all Tyr tac	non-onc
IgkS1-46	Val GTG	wt with all Val gtg	wt with all Val gtg	non-onc
IgkS1-47	Val GTA	wt with all Val gta	wt with all Val gta	non-onc
IgkS1-48	Val GTT	wt with all Val gtt	wt with all Val gtt	non-onc
IgkS1-49	Val GTC	wt with all Val gtc	wt with all Val gtc	non-onc
IgkS1-50	Leu CTG	altered with Leu ctg	altered with Leu ctg	onc
IgkS1-51	Leu CTA	altered with Leu cta	altered with Leu cta	onc
IgkS1-52	Leu CTT	altered with Leu ctt	altered with Leu ctt	onc
IgkS1-53	Leu CTC	altered with Leu ctc	altered with Leu ctc	onc
IgkS1-54	Leu TTG	altered with Leu ttg	altered with Leu ttg	onc
IgkS1-55	Leu TTA	altered with Leu tta	altered with Leu tta	onc
IgkS1-56	Pro CCG	altered with Pro ccg	altered with Pro ccg	onc
IgkS1-57	Pro CCA	altered with Pro cca	altered with Pro cca	onc
IgkS1-58	Pro CCT	altered with Pro cct	altered with Pro cct	onc
IgkS1-59	Pro CCC	altered with Pro ccc	altered with Pro ccc	onc
Secretory E7 construct series 2				
IgkS2-1	Ala GCG	mc	mc	linkerA-onc
IgkS2-2	Ala GCA	mc	mc	linkerA-onc
IgkS2-3	Ala GCT	mc	mc	linkerA-onc

Construct	AA & Codon	CU of Sec Seq	CU of E7	E7 Protein
IgkS2-4	Ala GCC	mc	mc	linkerA-onc
IgkS2-5	Arg AGG	mc	mc	linkerR-onc
IgkS2-6	Arg AGA	mc	mc	linkerR-onc
IgkS2-7	Arg CGG	mc	mc	linkerR-onc
IgkS2-8	Arg CGA	mc	mc	linkerR-onc
IgkS2-9	Arg CGT	mc	mc	linkerR-onc
IgkS2-10	Arg CGC	mc	mc	linkerR-onc
IgkS2-11	Asn AAT	mc	mc	linkerN-onc
IgkS2-12	Asn AAC	mc	mc	linkerN-onc
IgkS2-13	Asp GAT	wt with all Asp gat	wt with all Asp gat	onc
IgkS2-14	Asp GAC	wt with all Asp gac	wt with all Asp gac	onc
IgkS2-15	Cys TGT	wt	wt with all Cys tgt	onc
IgkS2-16	Cys TGC	wt	wt with all Cys tgc	onc
IgkS2-17	Glu GAG	wt with all Glu gag	wt with all Glu gag	onc
IgkS2-18	Glu GAA	wt with all Glu gaa	wt with all Glu gaa	onc
IgkS2-19	Gln CAG	wt	wt with all Gln cag	onc
IgkS2-20	Gln CAA	wt	wt with all Gln caa	onc
IgkS2-21	Gly GGG	wt with all Gly ggg	wt with all Gly ggg	onc
IgkS2-22	Gly GGA	wt with all Gly gga	wt with all Gly gga	onc
IgkS2-23	Gly GGT	wt with all Gly ggt	wt with all Gly ggt	onc
IgkS2-24	Gly GGC	wt with all Gly ggc	wt with all Gly ggc	onc

Construct	AA & Codon	CU of Sec Seq	CU of E7	E7 Protein
IgkS2-25	His CAT	mc	mc	linkerH-onc
IgkS2-26	His CAC	mc	mc	linkerH-onc
IgkS2-27	Ile ATA	wt	wt with all Ile ata	onc
IgkS2-28	Ile ATT	wt	wt with all Ile att	onc
IgkS2-29	Ile ATC	wt	wt with all Ile atc	onc
IgkS2-30	Lys AAG	mc	mc	linkerK-onc
IgkS2-31	Lys AAA	mc	mc	linkerK-onc
IgkS2-32	Phe TTT	mc	mc	linkerF-onc
IgkS2-33	Phe TTC	mc	mc	linkerF-onc
IgkS2-34	Ser AGT	wt with all Ser agt	wt with all Ser agt	onc
IgkS2-35	Ser AGC	wt with all Ser agc	wt with all Ser agc	onc
IgkS2-36	Ser TCG	wt with all Ser tcg	wt with all Ser tcg	onc
IgkS2-37	Ser TCA	wt with all Ser tca	wt with all Ser tca	onc
IgkS2-38	Ser TCT	wt with all Ser tct	wt with all Ser tct	onc
IgkS2-39	Ser TCC	wt	wt with all Ser tcc	onc
IgkS2-40	Thr ACG	wt with all Thr acg	wt with all Thr acg	onc
IgkS2-41	Thr ACA	wt with all Thr aca	wt with all Thr aca	onc
IgkS2-42	Thr ACT	wt with all Thr act	wt with all Thr act	onc
IgkS2-43	Thr ACC	wt with all Thr acc	wt with all Thr acc	onc
IgkS2-44	Tyr TAT	mc	mc	linkerY-onc
IgkS2-45	Tyr TAC	mc	mc	linkerY-onc

Construct	AA & Codon	CU of Sec Seq	CU of E7	E7 Protein
IgkS2-46	Val GTG	wt with all Val gtg	wt with all Val gtg	onc
IgkS2-47	Val GTA	wt with all Val gta	wt with all Val gta	onc
IgkS2-48	Val GTT	wt with all Val gtt	wt with all Val gtt	onc
IgkS2-49	Val GTC	wt with all Val gtc	wt with all Val gtc	onc
IgkS2-11b	Asn AAT	wt	wt with all Asn aat	linkerN-non-onc
IgkS2-12b	Asn AAC	wt	wt with all Asn aac	linkerN-non-onc

AA= amino acid, CU= codon usage, mc= mammalian consensus, wt= wild-type, onc= oncogenic, non-onc= non-oncogenic, Sec seq= secretory sequence, N/A= not applicable

Control constructs

[0173] Control E7 constructs were based on those from Liu *et al.* (2002). Both oncogenic (*i.e.* wild-type) and non-oncogenic E7 control constructs were made with wild-type or mammalian consensus codon usage. "Non-oncogenic" E7 is E7 with D21G, C24G, E26G mutations, *i.e.* with mutations that have been reported to render E7 non-transforming (Edmonds and Vousden, 1989; Heck et al, 1992).

[0174] The secretory sequence was derived from *Mus musculus* IgK RNA for the anti-HLA-DR antibody light chain (GenBank accession number D84070). For some constructs the codon usage of this sequence was modified.

Wild-type codon usage control constructs:

[0175] The wild-type (wt) codon usage E7 construct from Liu et al. was used as the template in a site-directed mutagenesis PCR to make the wt codon usage non-oncogenic E7 construct.

[0176] The non-oncogenic and oncogenic wild-type codon usage E7 sequences were amplified to incorporate a 5' *Bam*HI site and a 3' *Eco*RI site. The resulting fragments were cloned into *Bam*HI and *Eco*RI cut pCDNA3 and sequenced. The secretory fragment was

made by whole gene synthesis using wild-type codon usage with flanking *Kpn*I and *Bam*HI sites. The Kozak-secretory fragments were then ligated into *Kpn*I/*Bam*HI cut pCDNA3-wtE7 (non-oncogenic or oncogenic) to make **pCDNA3-Igk-nE7** and **pCDNA3-Igk-E7** (named **IgkC1 and IgkC3** respectively; see TABLE 12). The identity of the constructs was confirmed by sequencing.

Mammalian consensus (mc) codon usage control constructs:

[0177] As there were errors in the original mammalian consensus (mc) E7 construct (L28F, Q70R and an E35 deletion; Liu et al., 2002) it was not used. A mc non-oncogenic E7 control construct was synthesized by whole gene synthesis. A mc oncogenic E7 (*i.e.*, wild-type E7) control construct was subsequently made from the mc non-oncogenic E7 construct by single site-directed mutagenesis.

[0178] Secretory mc oncogenic and non-oncogenic constructs were made by amplifying the mc E7 sequence with a forward primer that introduced a *Bam*HI site and a reverse primer that incorporated an *Eco*RI site. The resulting E7 fragment was cloned into the respective sites in pCDNA3 and sequenced. A mc secretory sequence flanked by *Kpn*I and *Bam*HI sites, 5' and 3' respectively, was synthesised and ligated into the *Kpn*I and *Bam*HI sites of pCDNA3-mcE7 (oncogenic or non-oncogenic) to make **pCDNA3-mcIgk-mcnE7** and **pCDNA3-mcIgk-mcE7** (named **IgkC2 and IgkC4** respectively; see TABLE 12). The identity of the constructs was confirmed by sequencing.

Secreted non-oncogenic E7 constructs with predominantly wild-type codon usage, modified for individual codons

[0179] Plasmids encoding a non-oncogenic form of E7 were made for all of the codons, with the exception of the Pro and Leu codons, stop codons and codons for non-degenerate amino acids. As Phe occurs just once in the E7 sequence, the codons for two Leu residues, L15 and L22, were mutated to Phe codons. A combination of techniques was used to make these constructs. When few mutations were required single or multi site-directed mutagenesis of a control construct encoding non-oncogenic E7 was performed (details of the control construct are given above under "control constructs"). When more extensive modifications were required whole gene synthesis was employed. Regardless of the methods used these constructs all include an E7 encoding sequence with identical upstream and downstream sequence cloned into the *Kpn*I and *Eco*RI sites of pCDNA3. These constructs were then modified to include a secretory sequence, as described below.

[0180] First, using the whole gene synthesis method, DNA fragments that included a secretory sequence flanked by *Kpn*I and *Bam*HI sites were synthesized. For some constructs the amino acid of interest occurred in the secretory sequence so individual modified secretory sequence fragments were made. For constructs for amino acids that did not occur in the secretory sequence, wild-type secretory sequence was used. These fragments were digested with *Kpn*I and *Bam*HI. Then, using the relevant nE7 construct as a template and a standard PCR protocol, a *Bam*HI site was introduced at the 5' end of the E7 sequence. The 3' *Eco*RI site was retained. The resulting E7 fragments were cut with *Bam*HI and *Eco*RI, purified, and ligated into pCDNA3. Following sequencing, the plasmids were cut with *Kpn*I and *Bam*HI and ligated with the relevant *Kpn*I/*Bam*HI secretory sequences. The sequences of the constructs were then confirmed. Constructs **IgkS1-1 to IgkS1-49** were made in this way (see TABLE 12 and Figures 1 to 11, 13 and 15 to 17 for sequence comparisons).

Secreted E7 constructs with individual Pro or Leu codons modified

[0181] E7 DNA sequences in which the Pro or Leu codons were individually modified were designed. The rest of the codon usage for these E7 DNAs was the same for all of the Pro and Leu constructs but differed from the wild-type or mammalian consensus codon usage. [Note that this codon usage was based on our preliminary data from immunizing mice with the GFP constructs.]

[0182] The Pro/LeuE7 DNA fragments, flanked by *Hin*dIII and *Bam*HI sites, were made by whole gene synthesis and cloned into the *Hin*dIII and *Bam*HI sites of pCDNA3. Using these constructs as templates, a *Kpn*I site was incorporated upstream and an *Eco*RI site downstream, of the Pro/Leu E7 sequences by standard PCR methods. The resulting fragments were cut with *Kpn*I and *Eco*RI and cloned into pCDNA3. These constructs were then used to make the secreted E7 constructs with Pro or Leu codon modifications.

[0183] Firstly, using the whole gene synthesis method, DNA fragments that included a secretory sequence flanked by *Kpn*I and *Bam*HI sites were synthesized. As Pro and Leu occur in the secretory sequence, individually modified secretory sequence fragments were made for the different constructs. These fragments were digested with *Kpn*I and *Bam*HI. Then, using the relevant Pro or Leu E7 construct as a template and a standard PCR protocol, a *Bam*HI site was introduced at the 5' end of the E7 sequence. The 3' *Eco*RI site was retained. The resulting fragments were cut with *Bam*HI and *Eco*RI, purified, and ligated into pCDNA3. Following sequencing, the plasmids were cut with *Kpn*I and *Bam*HI and ligated with the

relevant *Kpn*I/ *Bam*HI secretory sequences. The resulting constructs were sequenced and are denoted **IgkS1-50 to IgkS1-59** (see TABLE 12 and Figures 12 and 14 for sequence comparisons).

Secreted E7 constructs with predominantly wild-type codon usage, modified for individual codons

[0184] Constructs encoding a secreted form of oncogenic E7 (*i.e.* wild-type E7 protein) were made by site-directed mutagenesis of the plasmids encoding a secreted form of non-oncogenic E7. This was done for constructs for codons for the following amino acids: Asp, Cys, Glu, Gln, Gly, Ile, Ser, Thr and Val.

[0185] Site-directed mutagenesis was carried out using the Quikchange II Site-directed Mutagenesis kit (Stratagene, La Jolla CA) and appropriate PAGE (polyacrylamide gel electrophoresis)-purified primers (Sigma) according to the manufacturer's instructions. The pCDNA-kIgkX-nE7X series of constructs were used as templates for the mutagenesis (*i.e.* constructs IgkS1-13 to 24, IgkS1-27 to 29, IgkS1-34 to 43 and IgkS1-46 to 49). The primers introduced the desired G21D, G24C, G26E mutations.

[0186] The resulting constructs, **IgkS2-13 to 24, IgkS2-27 to 29, IgkS2-34 to 43** and **IgkS2-46 to 49** (see Table 8, SEQ ID NOs: 1 to 29), have wild-type codon usage for the Igk secretory sequence and E7 sequence with the exception that the codons for the relevant amino acid were changed, and they encode oncogenic E7.

Linker constructs

[0187] Constructs encoding the N-terminal Igk secretory sequence followed by a linker sequence (XXGXGXX, where X is the relevant amino acid for a particular construct and G is glycine) and the E7 protein were made for each of the following amino acids: Asn, Ala, Lys, Arg, Phe, His and Tyr.

[0188] Fragments consisting of the Igk secretory sequence (with mammalian consensus codon usage) and the linker sequences were made by PCR using Taq polymerase and standard cycling conditions, as recommended by the manufacturer.

[0189] The fragments were amplified from pCDNA3-kmcIgk-mcE7 using a common forward primer (5'TTGAATAGGTACCGCCGCCACCATGGAGACCGACACCCTCC3'; SEQ IDNO: 90) that annealed to the *Kpn*I site, the Kozak sequence and the beginning of the Igk secretory

sequence. The reverse primers were different for each linker construct and annealed to the end of the Igk secretory sequence (with mammalian consensus codon usage), introduced new sequence that encoded the relevant linker sequence and a 3' *Bam*HI site.

[0190] The fragments were digested with *Kpn*I/*Bam*HI and were ligated into *Kpn*I/*Bam*HI-cut pCDNA3-mcIgk-mcE7 (*i.e.* the Kozak sequence and secretory sequence had been removed from the plasmid by digestion) to make pCDNA3-mcIgk-linkerX-mcE7 (*i.e.*, **IgkS2-1 to 12, IgkS2-25 and 26, IgkS2-30 to 33** and **IgkS2-44** and **45** as illustrated in Table 8, SEQ ID NOs: 30 to 49).

[0191] For Asn the fragments were also ligated into *Kpn*I/*Bam*HI-cut pCDNA3-Igk-nE7Asn1/2 (*i.e.* IgkS1-11 and 12) to make pCDNA3-mcIgk-linkerN1/2-nE7Asn1/2 (*i.e.*, **IgkS2-11b** and **IgkS2-12b**, see Table 12).

E7 Protein Expression

Cell culture

[0192] CHO cells were cultured in DMEM (GIBCO from Invitrogen) containing 10% foetal bovine serum (FBS) (DKSH), penicillin, streptomycin and glutamine (GIBCO from Invitrogen) at 37° C and 5% CO_2. Cells were plated into 6-well plates at 3 x 10^5 /well, 24 hours prior to transfection. For each transfection, 2 µg of DNA was mixed with 50 µL OptiMEM (GIBCO from Invitrogen) and 4 µL Plus reagent (Invitrogen) and incubated at room temperature (RT) for 30 min. Lipofectamine (Invitrogen; 5 µL in 50 µL OptiMEM) was added and the complexes incubated at RT for 30 min. The cells were rinsed with OptiMEM, 2 mL OptiMEM were added to each well, and the complexes then added. The cells were incubated overnight at 37°C and 5% CO_2. The following morning the complexes were removed and 2ml of fresh DMEM containing 2% FBS added to each well.

[0193] Cell pellets and supernatants were collected about 40 h after transfection. The cell pellets were resuspended in lysis buffer (0.1% NP-40, 2 µg/mL Aprotinin, 1 µg/mL Leupeptin and 2mM PMSF in PBS). Transfections were carried out in duplicate and repeated. Control transfections, with empty vector (pCDNA3), were also carried out.

Western blotting

[0194] Western blots of the CHO cell supernatants or lysates were carried out according to standard protocols. Briefly, this involved firstly separating the samples by polyacrylamide gel electrophoresis (PAGE). For cell lysates, 30 µg of total protein were

loaded for each sample. For supernatants, 30 µL of each was loaded. The protein samples were boiled with SDS-PAGE loading buffer for 10 mins before loading onto 12% SDS-PAGE gels and the gels were run at 150-200V for approximately 1h.

[0195] The separated proteins were then transferred from the gels to PVDF membrane (100V for 1h). The membranes were blocked with 5% skim milk (in PBS/0.05% Tween 20 (PBS-T)) for 1h at room temperature and were then incubated with the primary antibody, HPV-16 E7 Mouse Monoclonal Antibody (Zymed Laboratories) at a concentration of 1:1000 in 5% skim milk (in PBS-T) overnight at 4°C. Following washing of the membrane in PBS-T (3 X 10min), secondary antibody, anti-mouse IgG (Sigma) in 5% skim milk, was added and the membrane incubated at room temperature for 4h. The membranes were washed as before, incubated in a mixture containing equal volumes of solution A (4.425 mL water, 50 µL luminol, 22 µL p-coumaric and 500 µL 1M Tris pH 8.5) and solution B (4.5 mL water, 3 µL 30% H_2O_2 and 500 µL 1M Tris pH8.5) for 1 min, and then dried and wrapped in plastic wrap. Film was exposed to the blots for various times (1min, 3 min or 10 min) and the film then developed.

Gene Gun Immunization Protocols

Plasmid Purification

[0196] All plasmids used for vaccination were grown in the *Escherichia coli* strain DH5α and purified using the Nucleobond Maxi Kit (Machery-Nagal). DNA concentration was quantitated spectrophotometrically at 260 nm.

Preparation of DNA/gold cartridges

[0197] Coating of gold particles with plasmid DNA was performed as described in the Biorad Helios Gene Gun System instruction manual using a microcarrier loading quantity (MLQ) of 0.5mg gold/cartridge and a DNA loading ratio of 2 µg DNA/mg gold. This resulted in 1 µg of DNA per prepared cartridge. In brief 50 µL of 0.05M spermidine (Sigma) was added to 25 mg of 1.0 µm gold particles (Bio-Rad) and the spermidine/gold was sonicated for 3 seconds.. 50 µg of plasmid DNA was then added, followed by the dropwise addition of 100 µL 1M $CaCl_2$ while vortexing. The mixture was allowed to precipitate at room temperature for 10 min, then centrifuged to pellet the DNA/ gold. The pellet was washed three times with HPLC grade ethanol (Scharlau), before resuspension in HPLC grade ethanol containing 0.5 mg/mL of polyvinylpyrrolidone (PVP) (Bio-Rad). The gold/plasmid suspension was then coated onto Tefzel tubing and 0.5 inch cartridges prepared.

Gene Gun Immunization of Mice

[0198] Groups of 8 female C57BL6/J (6-8 weeks old) (ARC, WA or Monash Animal Services, VIC) were immunized on Day 0, Day21, Day 42 and Day 63 with the relevant DNA. The day before each immunization the abdomen of each mouse was shaved and depilatory cream (Nair) applied for 1 minute. DNA was delivered with the Helios gene gun (Biorad) using a pressure of 400psi. Mice were given 2 shots on either side of the abdomen, with 1 μg of DNA delivered per shot. Serum was collected via intra-ocular bleed 2 days prior to initial immunization and 2 weeks after each subsequent immunization (Day 2, Day 35, Day 56 and Day 77).

ELISA to measure E7 immune response

[0199] Nine peptides spanning the full-length of HPV16E7 (Frazer *et al.*, 1995) were used to measure the E7 antibody response. The peptides were synthesised and purified to >70% purity by Auspep (Melbourne). Peptides GF101 to 106 and GF108 to 109 described in Frazer *et al.* were made. Note that instead of GF107, GF107a was used: HYNIVTFCCKCDSTLRL.

[0200] GF102 D13G, GF103 D5G/C8G/E10G and GF104E2G peptides, named GF102n, GF103n and GF104n respectively, were also synthesised. These peptides were used for the ELISA when measuring antibodies to non-oncogenic E7 *i.e.* these peptides incorporate the mutations that were made to make the E7 protein non-oncogenic.

[0201] Microtiter plates were coated overnight with 50 μL of 10μg/mL E7 peptide per well. After coating, microtiter plates (Maxisorp, Nunc) were washed two times with PBS/0.05% Tween 20 (PBS-T) and then blocked for two hours at 37°C with 100 μL of 5% skim milk powder in PBS-T. After blocking, plates were washed three times with PBS-T and 50 μL of mouse sera at a dilution of 1 in 100 was added for 2 hours at 37°C. All serum was assayed in duplicate wells. Plates were then washed three times with PBS-T and 50 μL of sheep anti-mouse IgG horseradish peroxidise conjugate (Sigma) was added at a 1 in 1000 dilution. After 1 hour plates were washed and 50 μL of OPD substrate was added. Absorbance was measured after 30 min and the addition of 25 μL of 2.5 M HCl at 490 nm in a Multiskan EX plate reader (Pathtech). Note controls were included: control primary antibody for a positive control, secondary antibody only, and day 0 serum/ serum from unimmunized mice as negative controls.

[0202] The immune response preferences of codons determined from these experiments are tabulated in TABLE 1.

EXAMPLE 2

CONSTRUCTION OF CODON MODIFIED INFLUENZA A VIRUS (H5N1) HA DNA FOR CONFERRING AN ENHANCED IMMUNE RESPONSE TO H5N1 HA

[0203] The wild-type nucleotide sequence of the influenza A virus, HA gene for hemagglutinin (A/Hong Kong/213/03(H5N1), MDCK isolate, embryonated chicken egg isolate) is shown in SEQ ID NO: 50 and encodes the amino acid sequence shown in SEQ ID NO: 51. Several codons within that sequence were mutated using the method described in Example 1. Specifically, the method involved replacing codons of the wild type nucleotide sequence with corresponding synonymous codons having higher immune response preferences than the codons they replaced, as represented in Table 1. An illustrative codon modified nucleotide sequence comprising high immune response preference codons is shown in SEQ ID NO: 52.

EXAMPLE 3

CONSTRUCTION OF CODON MODIFIED INFLUENZA A VIRUS (H3N1) DNA FOR CONFERRING AN ENHANCED IMMUNE RESPONSE TO H3N1 HA

[0204] The wild-type nucleotide sequence of the influenza A virus, HA gene for hemagglutinin (A/swine/Korea/PZ72-1/2006(H3N1)) is shown in SEQ ID NO: 53 and encodes the amino acid sequence shown in SEQ ID NO: 54. Specifically, the method involved replacing codons of the wild type nucleotide sequence with corresponding synonymous codons having higher immune response preferences than the codons they replaced, as represented in Table 1. An illustrative codon modified nucleotide sequence comprising high immune response preference codons is shown in SEQ ID NO: 55.

EXAMPLE 4

CONSTRUCTION OF CODON MODIFIED INFLUENZA A VIRUS (H5N1) NA DNA FOR CONFERRING AN ENHANCED IMMUNE RESPONSE TO H5N1 NA

[0205] The wild-type nucleotide sequence of the influenza A virus, NA gene for neuraminidase (A/Hong Kong/213/03(H5N1), NA gene neuraminidase, MDCK isolate, embryonated chicken egg isolate) is shown in SEQ ID NO: 56 and encodes the amino acid

sequence shown in SEQ ID NO: 57. Several codons within that sequence were mutated using the method described in Example 1. Specifically, the method involved replacing codons of the wild type nucleotide sequence with corresponding synonymous codons having higher immune response preferences than the codons they replaced, as represented in Table 1. An illustrative codon modified nucleotide sequence comprising high immune response preference codons is shown in SEQ ID NO: 58.

EXAMPLE 5

CONSTRUCTION OF CODON MODIFIED INFLUENZA A VIRUS (H3N1) NA DNA FOR CONFERRING AN ENHANCED IMMUNE RESPONSE TO H3N1 NA

[0206] The wild-type nucleotide sequence of the influenza A virus, NA gene for neuraminidase (A/swine/MI/PU243/04(H3N1)) is shown in SEQ ID NO: 59 and encodes the amino acid sequence shown in SEQ ID NO: 60. Several codons within that sequence were mutated using the method described in Example 1. Specifically, the method involved replacing codons of the wild type nucleotide sequence with corresponding synonymous codons having higher immune response preferences than the codons they replaced, as represented in Table 1. An illustrative codon modified nucleotide sequence comprising high immune response preference codons is shown in SEQ ID NO: 61.

EXAMPLE 6

CONSTRUCTION OF CODON MODIFIED HEPATITIS C VIRUS E1 (1AH77) DNA FOR CONFERRING AN ENHANCED IMMUNE RESPONSE TO HCV E1 (1AH77)

[0207] The wild-type nucleotide sequence of the hepatitis C Virus E1, (serotype 1A, isolate H77, from polyprotein nucleotide sequence AF009606) is shown in SEQ ID NO: 62 and encodes the amino acid sequence (NP 751920) shown in SEQ ID NO: 63. Several codons within that sequence were mutated using the method described in Example 1. Specifically, the method involved replacing codons of the wild type nucleotide sequence with corresponding synonymous codons having higher immune response preferences than the codons they replaced, as represented in Table 1. An illustrative codon modified nucleotide sequence comprising high immune response preference codons is shown in SEQ ID NO: 64.

EXAMPLE 7

CONSTRUCTION OF CODON MODIFIED HEPATITIS C VIRUS E2 (1AH77) DNA FOR CONFERRING AN ENHANCED IMMUNE RESPONSE TO HCV E2 (1AH77)

[0208] The wild-type nucleotide sequence of the hepatitis C Virus E2, (serotype 1A, isolate H77, from polyprotein nucleotide sequence AF009606) is shown in SEQ ID NO: 65 and encodes the amino acid sequence (NP 751921) shown in SEQ ID NO: 66. Several codons within that sequence were mutated using the method described in Example 1. Specifically, the method involved replacing codons of the wild type nucleotide sequence with corresponding synonymous codons having higher immune response preferences than the codons they replaced, as represented in Table 1. An illustrative codon modified nucleotide sequence comprising high immune response preference codons is shown in in SEQ ID NO: 67.

EXAMPLE 8

CONSTRUCTION OF CODON MODIFIED EPSTEIN – BARR VIRUS TYPE 1 GP350 DNA FOR CONFERRING AN ENHANCED IMMUNE RESPONSE TO EBV TYPE 1 GP350

[0209] The wild-type nucleotide sequence of the Epstein - Barr virus, EBV type 1 gp350 (Gene BLLF1, strand 77142-79865) is shown in SEQ ID NO: 68 and encodes amino acid sequence (CAD53417) shown in SEQ ID NO: 69. Several codons within that sequence were mutated using the method described in Example 1.Specifically, the method involved replacing codons of the wild type nucleotide sequence with corresponding synonymous codons having higher immune response preferences than the codons they replaced, as represented in Table 1. An illustrative codon modified nucleotide sequence comprising high immune response preference codons is shown in SEQ ID NO: 70.

EXAMPLE 9

CONSTRUCTION OF CODON MODIFIED EPSTEIN – BARR VIRUS TYPE 2 GP350 DNA FOR CONFERRING AN ENHANCED IMMUNE RESPONSE TO EBV TYPE 2 GP350

[0210] The wild-type nucleotide sequence of the Epstein - Barr virus, EBV type 2 gp350 (Gene BLLF1, strand 77267-29936) is shown in SEQ ID NO: 71 and encodes the amino acid sequence (YP 001129462) shown in SEQ ID NO: 72. Several codons within that sequence were mutated using the method described in Example 1. Specifically, the method involved replacing codons of the wild type nucleotide sequence with corresponding

synonymous codons having higher immune response preferences than the codons they replaced, as represented in Table 1. An illustrative codon modified nucleotide sequence comprising high immune response preference codons is shown in SEQ ID NO: 73.

EXAMPLE 10

CONSTRUCTION OF CODON MODIFIED HERPES SIMPLEX VIRUS 2 GLYCOPROTEIN B DNA FOR CONFERRING AN ENHANCED IMMUNE RESPONSE TO HSV-2 GLYCOPROTEIN B

[0211] The wild-type nucleotide sequence of the Herpes Simplex virus 2, glycoprotein B strain HG52 (genome strain NC 001798) is shown in SEQ ID NO: 74 and encodes the amino acid sequence (CAB06752) shown in SEQ ID NO: 75. Several codons within that sequence were mutated using the method described in Example 1. Specifically, the method involved replacing codons of the wild type nucleotide sequence with corresponding synonymous codons having higher immune response preferences than the codons they replaced, as represented in Table 1. An illustrative codon modified nucleotide sequence comprising high immune response preference codons is shown in SEQ ID NO: 76.

EXAMPLE 11

CONSTRUCTION OF CODON MODIFIED HERPES SIMPLEX VIRUS 2 GLYCOPROTEIN D DNA FOR CONFERRING AN ENHANCED IMMUNE RESPONSE TO HSV-2 GLYCOPROTEIN D

[0212] The wild-type nucleotide sequence of the Herpes Simplex virus 2, glycoprotein D strain HG52 (genome strain NC 001798) is shown in SEQ ID NO: 77 and encodes the amino acid sequence (NP 044536) shown in SEQ ID NO: 78. Several codons within that sequence were mutated using the method described in Example 1. Specifically, the method involved replacing codons of the wild type nucleotide sequence with corresponding synonymous codons having higher immune response preferences than the codons they replaced, as represented in Table 1. An illustrative codon modified nucleotide sequence comprising high immune response preference codons is shown in SEQ ID NO: 79.

EXAMPLE 12

OPTIMISED E7 AND HSV-2 CONSTRUCTS

Design and synthesis of optimal and least optimal E7 constructs

[0213] One de-optimized (W) and three optimized (O1-O3) E7 constructs were designed and made using the codon preferences summarized in Table 1 ("the Immune

Coricode table"). The least favourable codons were used for construct W. For the first optimized construct, O1, whose sequence is shown in SEQ ID NO: 81, all of the codons were modified to those codons determined most optimal. O2, whose sequence is shown in SEQ ID NO: 82, is an alternative optimized construct which involved changing all Ala to GCT; Arg CGG and AGG to CGA and AGA, respectively; Glu to GAA; Gly to GGA; Ile to ATC; all Leu to CTG; Phe to TTT, Pro to CCT or CCC, Ser to TCG, Thr to ACG; and all Val except GTG to GTC. The O2 modifications avoided, with the exception of Leu and Ile, changing codons to mammalian consensus-preferred codons. For O3, whose sequence is shown in SEQ ID NO: 83, only certain amino acids for which particularly distinct differences were observed between codons, and for which the optimal codon(s) was not also a mammalian consensus preferred codon, were modified. In particular, in O3 all non-preferred Gly, Leu, Pro, Ser and Thr codons were changed to GGA, CTC, CCT, TCG and ACG, respectively, and where a preferred codon was already used it was not altered. Codons for other amino acids in O3 were not modified.

Humoral and cellular responses to biolistic immunization with the optimal and least optimal E7 constructs

[0214] As may be seen in Figure 18 (a) all three optimized constructs (O1 to O3) gave rise to significantly larger antibody responses than the wild-type construct as measured by both the peptide ELISA and a GST-E7 protein ELISA. The amplitudes of the response were not statistically different between the three optimized constructs. The de-optimized construct, W, whose sequence is shown in SEQ ID NO: 84, gave a very low antibody response, appearing slightly lower but not statistically different from the wild-type (wt) codon usage (CU) construct, whose sequence is shown in SEQ ID NO: 80. From the IFN-γ ELISPOT experiments, a representative example of which is shown in Figure 18, it appears that the codon preferences for maximizing the antibody response are similar to those required for maximising the T cell response: the de-optimized construct W failed to give a measurable response in the IFN-γ ELISPOT assay and two of the optimized constructs (O2 and O3) gave statistically significantly larger responses than the wild-type CU construct. Over the three repeats the responses to O2 and O3 were not statistically different from each other. Unexpectedly, and in contrast to the antibody trend, in two of the three repeat experiments O1 gave a similar cellular response to the wt CU construct, which was less than that achieved by the O2 or O3 constructs.

Humoral and cellular responses to immunization by intradermal injection with the optimal and least optimal E7 constructs

[0215] The humoral and cellular responses of mice to the optimized, wild-type CU and de-optimized constructs delivered by intradermal injection were also measured and the results are summarized in Figure 19. In general, similar trends were observed for intradermal injection as for biolistic delivery.

[0216] From the E7 protein ELISA, it is apparent that the three optimized constructs, O1-O3, were all significantly better at generating antibodies than the wild-type construct and that the de-optimized construct gave a very low antibody response similar to wild-type. The optimized constructs all gave rise to significantly more spots in the IFN-γ ELISPOT than the wild-type construct and the de-optimized construct failed to give rise to a measurable response.

[0217] The amplitudes of the antibody responses to gene gun immunization were larger than that for the intradermally (ID) delivered vaccines, despite the ID immunization delivering more than five times the dose.

Design and synthesis of optimal and least optimal HSV-2constructs

[0218] Three optimized (O1-O3; whose sequences are shown in SEQ ID NO: 86-88, respectively) and a de-optimized construct (W; whose sequence is shown in SEQ ID NO: 88) encoding full-length glycoprotein D from Herpes Simplex Virus 2 (gD2) were prepared. A control construct pCDNA3-gD2 with wt CU was also made. Wild-type CU, whose sequence is shown in SEQ ID NO: 85, is close to MC CU.

Humoral responses to biolistic and intradermal immunization with the optimal and least optimal gD2 constructs

[0219] C57Bl/6 mice were immunized in two groups (8 mice/construct; used intradermal injection (ID) and gene gun delivery) using the same immunization protocol as for the E7 constructs.

[0220] Group 1 included pCDNA3-gD2 and pCDNA3-gD2 O1. Group 2 included pCDNA3-gD2, pCDNA3-gD2 O2, pCDNA3-gD2 O3, and pCDNA3-gD2 W.

[0221] Antibody responses were measured by an ELISA using plates coated with CHO cell supernatant containing C-terminally His tagged and truncated gD2. The truncation is at amino acid residue 331 and removes the transmembrane region resulting in the protein

being secreted into the medium. Control ELISA plates coated with supernatant from CHO cells transfected with empty vector were used as a control.

[0222] For both biolistic and intradermal injection delivery routes it was found that the three optimized constructs generated similar levels of antibodies as the wt CU gD2 construct (Figure 20). The de-optimized construct, W gD2, was very poor at generating antibodies, particularly when delivered by intradermal injection. The two delivery methods resulted in similar levels of antibodies.

[0223] To date, there are no DNA vaccines on the market for the treatment or prevention of disease in humans. There is a need to maximize the immune responses generated by DNA vaccines and the present invention discloses ways of enhancing efficacy of DNA vaccines by using codons that have a higher preference for producing an immune response.

[0224] The study described in this Example has validated the Immune Coricode table by applying it to optimization or de-optimization of the HPV16 E7 and HSV-2 glycoprotein D (gD2) genes and demonstrating that this does enhance or reduce, respectively, the antibody or cellular response to biolistic delivery of these genes to mammals such as mice.

Material and Methods

ELISPOT Assay

[0225] For the IFN-γ ELISPOTs, mice were immunized twice, at days 0 and 21, and the spleens were collected 3 weeks after the second immunization.

Intradermal Injection Protocol

[0226] The timing and frequency of the immunizations by intradermal injection were the same as for gene gun immunization. At each immunization 5 μg of DNA was injected per ear *i.e.* a total of 10 μg was administered per immunization per mouse. Hair removal prior to immunization was not necessary. The timing of bleeds and spleen collection was the same as for the gene gun immunized mice.

GST-E7 ELISA

[0227] The GST-E7 ELISA was carried out in the same way as the peptide ELISA with the exception that the plates were coated overnight with 50 μL of 10 μg/mL GST-tagged

E7 protein (kindly provided by the Frazer group from the Diamantina Institute, The University of Queensland, Brisbane).

HSV-2 gD ELISA

[0228] This ELISA was carried out in the same way as the E7 ELISAs with the exception that the plates were coated with supernatant from CHO cells transfected with a vector encoding C-terminally His-tagged and truncated gD2 protein. Control plates coated with supernatant from CHO cells transfected with empty vector were also used.

Detection of HPV-specific responses

[0229] For the detection of HPV-specific responses, 96-well filter ELISPOT plates (Millipore) were coated overnight with 10 μg/mL HPV GST-tagged E7 protein in 0.1 M $NaHCO_3$. For the detection of total IgG secreting cells, 96-well filter ELISPOT plates were coated overnight with 2 μg/mL goat anti-mouse Ig (Sigma) in PBS without $MgCl_2$ and $CaCl_2$. After coating, plates were washed once with complete DMEM without FCS and then blocked with complete DMEM supplemented with 10% FCS for one hour at 37° C. Cultured mouse spleen cells were washed and added to ELISPOT plates at 10^6 cells/100 μL. For the detection of HPV-specific memory B cells, plates were incubated overnight at 37° C and for measuring total IgG cells, plates were incubated for 1 hour at 37° C. For detection, we used biotinylated goat anti-mouse IgG (Sigma) in PBS-T/1% FCS, followed by 5 μg/mL HRP-conjugated avidin (Pierce) and developed using 3-amino-9-ethylcarbozole (Sigma). Developed plates were counted using an automated ELISPOT plate counter.

E7 IFN-γ ELISPOT

[0230] 96-well filter plates (Millipore) were coated overnight with 4 μg/mL of monoclonal antibody (AN18; Mabtech). After coating, plates were washed once with complete RPMI and blocked for 2 hours with complete RPMI with 10% foetal calf serum (FCS; CSL Ltd). Mouse spleens were made into single cell suspensions and treated with ACK lysis buffer, washed and resuspended at a concentration of 10^7 cells/mL. Spleen cells (10^6/well) were added to each well followed by the addition of complete RPMI supplemented with recombinant hIL-2 (ProSpec-Tany TechnoGene Ltd) and peptide to a final concentration of 10IU/well and 1 μg/mL, respectively. Medium containing hIL-2 without peptide was added to control wells. Plates were incubated for approximately 18 hours at 37° C in 5-8% CO_2.

[0231] After overnight incubation, cells were lysed by rinsing the plates in tap water and then washed six times in PBS/0.05% Tween 20 (PBS-T). For detection, biotinylated detection mAb (R4-6A2; Mabtech) in PBS-T/2% FCS was added, followed by horse radish peroxidase (HRP)-conjugated strepavidin and DAB (Sigma). Developed plates were counted using an automated ELISPOT plate counter.

[0232] The disclosure of every patent, patent application, and publication cited herein is hereby incorporated herein by reference in its entirety.

[0233] The citation of any reference herein should not be construed as an admission that such reference is available as "Prior Art" to the instant application.

[0234] Throughout the specification the aim has been to describe the preferred embodiments of the invention without limiting the invention to any one embodiment or specific collection of features. Those of skill in the art will therefore appreciate that, in light of the instant disclosure, various modifications and changes can be made in the particular embodiments exemplified without departing from the scope of the present invention. All such modifications and changes are intended to be included within the scope of the appended claims.

BIBLIOGRAPHY

[0235] Ausubel, F.M. (Ed.) 2007. Current Protocols in Molecular Biology. Ebook (http://www.mrw.interscience.wiley.com/emrw/9780471142720/cp/cpmb/toc).

[0236] Edmonds, C., and Vousden, K.H. (1989). A point mutational analysis of human papillomavirus type 16 E7 protein. Journal of Virology. 63: 2650-2656.

[0237] Frazer, I.H., Leippe, D.M., Dunn, L.A., Leim, A., Tindle, R.W., Fernando, G.J., Phelps, W.C., and Lambert, P.F. (1995). Immunological responses in human papillomavirus 16 E6/E7 transgenic mice to E7 protein correlate with the presence of skin disease. Cancer Research. 55: 2635-2639.

[0238] Heck, D.V., Yee, C.L., Howley, P.M., and Munger, K. (1992). Efficiency of binding the retinoblastoma protein correlates with the transforming capacity of the E7 oncoproteins of the human papillomaviruses. PNAS 89: 4442-4446.

[0239] Liu, W.J., Gao, F., Zhao, KN., Zhao, W., Fernando, G.J, Thomas, R. And Frazer, I.H. (2002). Codon modified human papillomavirus type 16 E7 DNA vaccine enhances cytotoxic T-lymphocyte induction and anti-tumour activity. Virology 301: 43-52.

[0240] Smith, H.O., Hutchison III, C.A., Pfannkoch, C. and Venter, J.C. (2003). Generating a synthetic genome by whole genome assembly: ϕX174 bacteriophage from synthetic oligonucleotides. PNAS. 100 (26): 15440-15445.

WHAT IS CLAIMED IS:

1. A method for constructing a synthetic polynucleotide from which a polypeptide is producible to confer an immune response to a target antigen in a mammal in a different quality than that conferred by a parent polynucleotide that encodes the same polypeptide, wherein the polypeptide corresponds to at least a portion of the target antigen, the method comprising: (a) selecting a first codon of the parent polynucleotide for replacement with a synonymous codon, wherein the synonymous codon is selected on the basis that it exhibits a different preference for conferring an immune response ("an immune response preference") than the first codon in a comparison of immune response preferences; and (b) replacing the first codon with the synonymous codon to construct the synthetic polynucleotide, wherein the comparison of immune response preferences of the codons is represented by TABLE 1:

TABLE 1

Amino Acid	Ranking of Immune Response Preferences for Synonymous Codons
Ala	$Ala^{GCT} > Ala^{GCC} > (Ala^{GCA}, Ala^{GCG})$
Arg	$(Arg^{CGA}, Arg^{CGC}, Arg^{CGT}, Arg^{AGA}) > (Arg^{AGG}, Arg^{CGG})$
Asn	$Asn^{AAC} > Asn^{AAT}$
Asp	$Asp^{GAC} > Asp^{GAT}$
Cys	$Cys^{TGC} > Cys^{TGT}$
Glu	$Glu^{GAA} > Glu^{GAG}$
Gln	$Gln^{CAA} = Gln^{CAG}$
Gly	$Gly^{GGA} > (Gly^{GGG}, Gly^{GGT}, Gly^{GGC})$
His	$His^{CAC} = His^{CAT}$
Ile	$Ile^{ATC} >> Ile^{ATT} > Ile^{ATA}$
Leu	$(Leu^{CTG}, Leu^{CTC}) > (Leu^{CTA}, Leu^{CTT}) >> Leu^{TTG} > Leu^{TTA}$
Lys	$Lys^{AAG} = Lys^{AAA}$
Phe	$Phe^{TTT} > Phe^{TTC}$

Pro	$Pro^{CCC} > Pro^{CCT} >> (Pro^{CCA}, Pro^{CCG})$
Ser	$Ser^{TCG} >> (Ser^{TCT}, Ser^{TCA}, Ser^{TCC}) >> (Ser^{AGC}, Ser^{AGT})$
Thr	$Thr^{ACG} > Thr^{ACC} >> Thr^{ACA} > Thr^{ACT}$
Tyr	$Tyr^{TAC} > Tyr^{TAT}$
Val	$(Val^{GTG}, Val^{GTC}) > Val^{GTT} > Val^{GTA}$

2. The method according to claim 1, wherein the synthetic polynucleotide confers a stronger or enhanced immune response than the parent polynucleotide under the same conditions and wherein the first and synonymous codons are selected from TABLE 2:

TABLE 2

First Codon	Synonymous Codon	First Codon	Synonymous Codon	First Codon	Synonymous Codon
Ala^{GCG}	Ala^{GCT}	Ile^{ATA}	Ile^{ATC}	Ser^{AGT}	Ser^{TCG}
Ala^{GCG}	Ala^{GCC}	Ile^{ATA}	Ile^{ATT}	Ser^{AGT}	Ser^{TCT}
Ala^{GCA}	Ala^{GCT}	Ile^{ATT}	Ile^{ATC}	Ser^{AGT}	Ser^{TCA}
Ala^{GCA}	Ala^{GCC}			Ser^{AGT}	Ser^{TCC}
Ala^{GCC}	Ala^{GCT}	Leu^{TTA}	Leu^{CTG}	Ser^{AGC}	Ser^{TCG}
		Leu^{TTA}	Leu^{CTC}	Ser^{AGC}	Ser^{TCT}
Arg^{CGG}	Arg^{CGA}	Leu^{TTA}	Leu^{CTA}	Ser^{AGC}	Ser^{TCA}
Arg^{CGG}	Arg^{CGC}	Leu^{TTA}	Leu^{CTT}	Ser^{AGC}	Ser^{TCC}
Arg^{CGG}	Arg^{CGT}	Leu^{TTA}	Leu^{TTG}	Ser^{TCC}	Ser^{TCG}
Arg^{CGG}	Arg^{AGA}	Leu^{TTG}	Leu^{CTG}	Ser^{TCA}	Ser^{TCG}
Arg^{AGG}	Arg^{CGA}	Leu^{TTG}	Leu^{CTC}	Ser^{TCT}	Ser^{TCG}
Arg^{AGG}	Arg^{CGC}	Leu^{TTG}	Leu^{CTA}		
Arg^{AGG}	Arg^{CGT}	Leu^{TTG}	Leu^{CTT}	Thr^{ACT}	Thr^{ACG}

Arg^{AGG}	Arg^{AGA}	Leu^{CTT}	Leu^{CTG}	Thr^{ACT}	Thr^{ACC}
		Leu^{CTT}	Leu^{CTC}	Thr^{ACT}	Thr^{ACA}
Asn^{AAT}	Asn^{AAC}	Leu^{CTA}	Leu^{CTG}	Thr^{ACA}	Thr^{ACG}
		Leu^{CTA}	Leu^{CTC}	Thr^{ACA}	Thr^{ACC}
Asp^{GAT}	Asp^{GAC}			Thr^{ACC}	Thr^{ACG}
		Phe^{TTC}	Phe^{TTT}		
Cys^{TGT}	Cys^{TGC}			Tyr^{TAT}	Tyr^{TAC}
		Pro^{CCG}	Pro^{CCC}		
Glu^{GAG}	Glu^{GAA}	Pro^{CCG}	Pro^{CCT}	Val^{GTA}	Val^{GTG}
		Pro^{CCA}	Pro^{CCC}	Val^{GTA}	Val^{GTC}
Gly^{GGC}	Gly^{GGA}	Pro^{CCA}	Pro^{CCT}	Val^{GTA}	Val^{GTT}
Gly^{GGT}	Gly^{GGA}	Pro^{CCT}	Pro^{CCC}	Val^{GTT}	Val^{GTG}
Gly^{GGG}	Gly^{GGA}			Val^{GTT}	Val^{GTC}

3. The method according to claim 1, wherein the synthetic polynucleotide confers a stronger or enhanced immune response than the parent polynucleotide under the same conditions and wherein the first and synonymous codons are selected from TABLE 3:

TABLE 3

First Codon	Synonymous Codon	First Codon	Synonymous Codon	First Codon	Synonymous Codon
Ala^{GCG}	Ala^{GCT}	Leu^{TTA}	Leu^{CTA}	Ser^{AGT}	Ser^{TCG}
Ala^{GCA}	Ala^{GCT}	Leu^{TTA}	Leu^{CTT}	Ser^{AGT}	Ser^{TCT}
Ala^{GCC}	Ala^{GCT}	Leu^{TTA}	Leu^{TTG}	Ser^{AGT}	Ser^{TCA}
		Leu^{TTG}	Leu^{CTA}	Ser^{AGC}	Ser^{TCG}
Arg^{CGG}	Arg^{CGA}	Leu^{TTG}	Leu^{CTT}	Ser^{AGC}	Ser^{TCT}

Arg^{CGG}	Arg^{CGT}			Ser^{AGC}	Ser^{TCA}
Arg^{CGG}	Arg^{AGA}	Phe^{TTC}	Phe^{TTT}	Ser^{AGC}	Ser^{TCC}
Arg^{AGG}	Arg^{CGA}			Ser^{TCC}	Ser^{TCG}
Arg^{AGG}	Arg^{CGT}	Pro^{CCG}	Pro^{CCT}	Ser^{TCA}	Ser^{TCG}
Arg^{AGG}	Arg^{AGA}	Pro^{CCA}	Pro^{CCT}	Ser^{TCT}	Ser^{TCG}
Glu^{GAG}	Glu^{GAA}			Thr^{ACT}	Thr^{ACG}
				Thr^{ACT}	Thr^{ACA}
Gly^{GGC}	Gly^{GGA}			Thr^{ACA}	Thr^{ACG}
Gly^{GGT}	Gly^{GGA}			Thr^{ACC}	Thr^{ACG}
Gly^{GGG}	Gly^{GGA}				
				Val^{GTA}	Val^{GTT}

4. The method according to claim 3, further comprising selecting a second codon of the parent polynucleotide for replacement with a synonymous codon, wherein the synonymous codon is selected on the basis that it exhibits a higher immune response preference than the second codon in a comparison of immune response preferences; and (b) replacing the second codon with the synonymous codon, wherein the comparison of immune response preferences of the codons is represented by TABLE 4:

TABLE 4

Second Codon	Synonymous Codon	Second Codon	Synonymous Codon	Second Codon	Synonymous Codon
Ala^{GCG}	Ala^{GCT}	Ile^{ATA}	Ile^{ATC}	Ser^{AGT}	Ser^{TCG}
Ala^{GCG}	Ala^{GCC}	Ile^{ATA}	Ile^{ATT}	Ser^{AGT}	Ser^{TCT}
Ala^{GCA}	Ala^{GCT}	Ile^{ATT}	Ile^{ATC}	Ser^{AGT}	Ser^{TCA}
Ala^{GCA}	Ala^{GCC}			Ser^{AGT}	Ser^{TCC}

Ala^{GCC}	Ala^{GCT}	Leu^{TTA}	Leu^{CTG}	Ser^{AGC}	Ser^{TCG}
		Leu^{TTA}	Leu^{CTC}	Ser^{AGC}	Ser^{TCT}
Arg^{CGG}	Arg^{CGA}	Leu^{TTA}	Leu^{CTA}	Ser^{AGC}	Ser^{TCA}
Arg^{CGG}	Arg^{CGC}	Leu^{TTA}	Leu^{CTT}	Ser^{AGC}	Ser^{TCC}
Arg^{CGG}	Arg^{CGT}	Leu^{TTA}	Leu^{TTG}	Ser^{TCC}	Ser^{TCG}
Arg^{CGG}	Arg^{AGA}	Leu^{TTG}	Leu^{CTG}	Ser^{TCA}	Ser^{TCG}
Arg^{AGG}	Arg^{CGA}	Leu^{TTG}	Leu^{CTC}	Ser^{TCT}	Ser^{TCG}
Arg^{AGG}	Arg^{CGC}	Leu^{TTG}	Leu^{CTA}		
Arg^{AGG}	Arg^{CGT}	Leu^{TTG}	Leu^{CTT}	Thr^{ACT}	Thr^{ACG}
Arg^{AGG}	Arg^{AGA}	Leu^{CTT}	Leu^{CTG}	Thr^{ACT}	Thr^{ACC}
		Leu^{CTT}	Leu^{CTC}	Thr^{ACT}	Thr^{ACA}
Asn^{AAT}	Asn^{AAC}	Leu^{CTA}	Leu^{CTG}	Thr^{ACA}	Thr^{ACG}
		Leu^{CTA}	Leu^{CTC}	Thr^{ACA}	Thr^{ACC}
Asp^{GAT}	Asp^{GAC}			Thr^{ACC}	Thr^{ACG}
		Phe^{TTC}	Phe^{TTT}		
Cys^{TGT}	Cys^{TGC}			Tyr^{TAT}	Tyr^{TAC}
		Pro^{CCG}	Pro^{CCC}		
Glu^{GAG}	Glu^{GAA}	Pro^{CCG}	Pro^{CCT}	Val^{GTA}	Val^{GTG}
		Pro^{CCA}	Pro^{CCC}	Val^{GTA}	Val^{GTC}
Gly^{GGC}	Gly^{GGA}	Pro^{CCA}	Pro^{CCT}	Val^{GTA}	Val^{GTT}
Gly^{GGT}	Gly^{GGA}	Pro^{CCT}	Pro^{CCC}	Val^{GTT}	Val^{GTG}
Gly^{GGG}	Gly^{GGA}			Val^{GTT}	Val^{GTC}

5. The method according to claim 1, wherein the synthetic polynucleotide confers a weaker or reduced immune response than the parent polynucleotide under the

same conditions and wherein the first and synonymous codons are selected from the TABLE 5:

TABLE 5

First Codon	Synonymous Codon	First Codon	Synonymous Codon	First Codon	Synonymous Codon
Ala^{GCT}	Ala^{GCG}	Ile^{ATC}	Ile^{ATA}	Ser^{TCG}	Ser^{TCT}
Ala^{GCT}	Ala^{GCA}	Ile^{ATC}	Ile^{ATT}	Ser^{TCG}	Ser^{TCA}
Ala^{GCT}	Ala^{GCC}	Ile^{ATT}	Ile^{ATA}	Ser^{TCG}	Ser^{TCC}
Ala^{GCC}	Ala^{GCG}			Ser^{TCG}	Ser^{AGC}
Ala^{GCC}	Ala^{GCA}	Leu^{CTG}	Leu^{CTA}	Ser^{TCG}	Ser^{AGT}
		Leu^{CTG}	Leu^{CTT}	Ser^{TCT}	Ser^{AGC}
Arg^{CGA}	Arg^{AGG}	Leu^{CTG}	Leu^{TTG}	Ser^{TCT}	Ser^{AGT}
Arg^{CGA}	Arg^{CGG}	Leu^{CTG}	Leu^{TTA}	Ser^{TCA}	Ser^{AGC}
Arg^{CGC}	Arg^{AGG}	Leu^{CTC}	Leu^{CTA}	Ser^{TCA}	Ser^{AGT}
Arg^{CGC}	Arg^{CGG}	Leu^{CTC}	Leu^{CTT}	Ser^{TCC}	Ser^{AGC}
Arg^{CGT}	Arg^{AGG}	Leu^{CTC}	Leu^{TTG}	Ser^{TCC}	Ser^{AGT}
Arg^{CGT}	Arg^{CGG}	Leu^{CTC}	Leu^{TTA}		
Arg^{AGA}	Arg^{AGG}	Leu^{CTA}	Leu^{TTG}	Thr^{ACG}	Thr^{ACC}
Arg^{AGA}	Arg^{CGG}	Leu^{CTA}	Leu^{TTA}	Thr^{ACG}	Thr^{ACA}
		Leu^{CTT}	Leu^{TTG}	Thr^{ACG}	Thr^{ACT}
Asn^{AAC}	Asn^{AAT}	Leu^{CTT}	Leu^{TTA}	Thr^{ACC}	Thr^{ACA}
		Leu^{TTG}	Leu^{TTA}	Thr^{ACC}	Thr^{ACT}
Asp^{GAC}	Asp^{GAT}			Thr^{ACA}	Thr^{ACT}
		Phe^{TTT}	Phe^{TTC}		
Cys^{TGC}	Cys^{TGT}			Tyr^{TAC}	Tyr^{TAT}
		Pro^{CCC}	Pro^{CCT}		

Glu^{GAA}	Glu^{GAG}	Pro^{CCC}	Pro^{CCA}	Val^{GTG}	Val^{GTT}
		Pro^{CCC}	Pro^{CCG}	Val^{GTG}	Val^{GTA}
Gly^{GGA}	Gly^{GGC}	Pro^{CCT}	Pro^{CCA}	Val^{GTC}	Val^{GTT}
Gly^{GGA}	Gly^{GGT}	Pro^{CCT}	Pro^{CCG}	Val^{GTC}	Val^{GTA}
Gly^{GGA}	Gly^{GGG}			Val^{GTT}	Val^{GTA}

6. The method according to claim 1, wherein the synthetic polynucleotide confers a weaker or reduced immune response than the parent polynucleotide under the same conditions and wherein the first and synonymous codons are selected from TABLE 6:

TABLE 6

First Codon	Synonymous Codon	First Codon	Synonymous Codon	First Codon	Synonymous Codon
Ala^{GCT}	Ala^{GCG}	Leu^{CTA}	Leu^{TTG}	Ser^{TCG}	Ser^{TCT}
Ala^{GCT}	Ala^{GCA}	Leu^{CTA}	Leu^{TTA}	Ser^{TCG}	Ser^{TCA}
Ala^{GCT}	Ala^{GCC}	Leu^{CTT}	Leu^{TTG}	Ser^{TCG}	Ser^{TCC}
		Leu^{CTT}	Leu^{TTA}	Ser^{TCG}	Ser^{AGC}
Arg^{CGA}	Arg^{AGG}	Leu^{TTG}	Leu^{TTA}	Ser^{TCG}	Ser^{AGT}
Arg^{CGA}	Arg^{CGG}			Ser^{TCT}	Ser^{AGC}
Arg^{CGT}	Arg^{AGG}	Phe^{TTT}	Phe^{TTC}	Ser^{TCT}	Ser^{AGT}
Arg^{CGT}	Arg^{CGG}			Ser^{TCA}	Ser^{AGC}
Arg^{AGA}	Arg^{AGG}	Pro^{CCT}	Pro^{CCA}	Ser^{TCA}	Ser^{AGT}
Arg^{AGA}	Arg^{CGG}	Pro^{CCT}	Pro^{CCG}	Ser^{TCC}	Ser^{AGC}
Glu^{GAA}	Glu^{GAG}			Thr^{ACG}	Thr^{ACC}
				Thr^{ACG}	Thr^{ACA}

First Codon	Synonymous Codon	First Codon	Synonymous Codon	First Codon	Synonymous Codon
Gly^{GGA}	Gly^{GGC}			Thr^{ACG}	Thr^{ACT}
Gly^{GGA}	Gly^{GGT}			Thr^{ACA}	Thr^{ACT}
Gly^{GGA}	Gly^{GGG}				
				Val^{GTT}	Val^{GTA}

7. The method according to claim 6, further comprising selecting a second codon of the parent polynucleotide for replacement with a synonymous codon, wherein the synonymous codon is selected on the basis that it exhibits a lower immune response preference than the second codon in a comparison of immune response preferences; and; (b) replacing the second codon with the synonymous codon, wherein the comparison of immune response preferences of the codons is represented by TABLE 7:

TABLE 7

Second Codon	Synonymous Codon	Second Codon	Synonymous Codon	Second Codon	Synonymous Codon
Ala^{GCT}	Ala^{GCG}	Ile^{ATC}	Ile^{ATA}	Ser^{TCG}	Ser^{TCT}
Ala^{GCT}	Ala^{GCA}	Ile^{ATC}	Ile^{ATT}	Ser^{TCG}	Ser^{TCA}
Ala^{GCT}	Ala^{GCC}	Ile^{ATT}	Ile^{ATA}	Ser^{TCG}	Ser^{TCC}
Ala^{GCC}	Ala^{GCG}			Ser^{TCG}	Ser^{AGC}
Ala^{GCC}	Ala^{GCA}	Leu^{CTG}	Leu^{CTA}	Ser^{TCG}	Ser^{AGT}
		Leu^{CTG}	Leu^{CTT}	Ser^{TCT}	Ser^{AGC}
Arg^{CGA}	Arg^{AGG}	Leu^{CTG}	Leu^{TTG}	Ser^{TCT}	Ser^{AGT}
Arg^{CGA}	Arg^{CGG}	Leu^{CTG}	Leu^{TTA}	Ser^{TCA}	Ser^{AGC}
Arg^{CGC}	Arg^{AGG}	Leu^{CTC}	Leu^{CTA}	Ser^{TCA}	Ser^{AGT}
Arg^{CGC}	Arg^{CGG}	Leu^{CTC}	Leu^{CTT}	Ser^{TCC}	Ser^{AGC}

Arg^{CGT}	Arg^{AGG}	Leu^{CTC}	Leu^{TTG}	Ser^{TCC}	Ser^{AGT}
Arg^{CGT}	Arg^{CGG}	Leu^{CTC}	Leu^{TTA}		
Arg^{AGA}	Arg^{AGG}	Leu^{CTA}	Leu^{TTG}	Thr^{ACG}	Thr^{ACC}
Arg^{AGA}	Arg^{CGG}	Leu^{CTA}	Leu^{TTA}	Thr^{ACG}	Thr^{ACA}
		Leu^{CTT}	Leu^{TTG}	Thr^{ACG}	Thr^{ACT}
Asn^{AAC}	Asn^{AAT}	Leu^{CTT}	Leu^{TTA}	Thr^{ACC}	Thr^{ACA}
		Leu^{TTG}	Leu^{TTA}	Thr^{ACC}	Thr^{ACT}
Asp^{GAC}	Asp^{GAT}			Thr^{ACA}	Thr^{ACT}
		Phe^{TTT}	Phe^{TTC}		
Cys^{TGC}	Cys^{TGT}			Tyr^{TAC}	Tyr^{TAT}
		Pro^{CCC}	Pro^{CCT}		
Glu^{GAA}	Glu^{GAG}	Pro^{CCC}	Pro^{CCA}	Val^{GTG}	Val^{GTT}
		Pro^{CCC}	Pro^{CCG}	Val^{GTG}	Val^{GTA}
Gly^{GGA}	Gly^{GGC}	Pro^{CCT}	Pro^{CCA}	Val^{GTC}	Val^{GTT}
Gly^{GGA}	Gly^{GGT}	Pro^{CCT}	Pro^{CCG}	Val^{GTC}	Val^{GTA}
Gly^{GGA}	Gly^{GGG}			Val^{GTT}	Val^{GTA}

8. A synthetic polynucleotide constructed according to any one of claims 1 to 7.

9. A chimeric construct that comprises a synthetic polynucleotide constructed according to claim 8, wherein the synthetic polynucleotide is operably connected to a regulatory polynucleotide.

10. A pharmaceutical composition that is useful for modulating an immune response to a target antigen in a mammal, which response is conferred by the expression of a parent polynucleotide that encodes a polypeptide corresponding to at least a portion of the target antigen, the composition comprising a chimeric construct and a pharmaceutically acceptable excipient and/or carrier, wherein the chimeric construct

comprises a synthetic polynucleotide that is operably connected to a regulatory polynucleotide and that is distinguished from the parent polynucleotide by the replacement of a first codon in the parent polynucleotide with a synonymous codon that has a different immune response preference than the first codon and wherein the first and synonymous codons are selected according to any one of TABLES 2, 3, 5 and 6.

11. The composition according to claim 10, further comprising an adjuvant that enhances the effectiveness of the immune response.

12. The composition according to claim 10, which is formulated for transcutaneous administration.

13. The composition according to claim 10, which is formulated for epidermal administration.

14. The composition according to claim 10, which is formulated for dermal administration.

15. The composition according to claim 10, which is formulated for intradermal administration.

16. The composition according to claim 10, which is formulated for biolistic delivery.

17. The composition according to claim 10, which is formulated for microneedle delivery.

18. The composition according to claim 10, which is formulated for intradermal injection.

19. The composition according to claim 10, wherein the synthetic polynucleotide confers a stronger or enhanced immune response than the parent polynucleotide under the same conditions and wherein the first and synonymous codons are selected according to TABLES 2 or 3.

20. The composition according to claim 10, wherein the synthetic polynucleotide confers a weaker or reduced immune response than the parent polynucleotide under the same conditions and wherein the first and synonymous codons are selected according to TABLES 5 or 6.

21. A method of modulating the quality of an immune response to a target antigen in a mammal, which response is conferred by the expression of a parent polynucleotide that encodes a polypeptide corresponding to at least a portion of the target antigen, the method comprising: introducing into the mammal a synthetic

polynucleotide that is operably connected to a regulatory polynucleotide and that is distinguished from the parent polynucleotide by the replacement of a first codon in the parent polynucleotide with a synonymous codon that has a different immune response preference than the first codon and wherein the first and synonymous codons are selected according to any one of TABLES 2, 3, 5 and 6.

22. The method according to claim 21, wherein expression of the synthetic polynucleotide results in a different quality of immune response than the one obtained through expression of the parent polynucleotide under the same conditions.

23. The method according to claim 21, wherein the chimeric construct is introduced into the mammal by delivering the construct to antigen-presenting cells of the mammal.

24. The method according to claim 21, wherein the chimeric construct is introduced into the dermis and/or epidermis of the mammal.

25. The method according to claim 21, wherein the chimeric construct is introduced into the mammal by transcutaneous administration.

26. The method according to claim 21, wherein the chimeric construct is introduced into the mammal by epidermal administration.

27. The method according to claim 21, wherein the chimeric construct is introduced into the mammal by intradermal administration.

28. The method according to claim 21, wherein the chimeric construct is introduced into the dermis or epidermis of the abdomen.

29. The method according to claim 21, wherein the immune response is selected from a cell-mediated response and a humoral immune response.

30. The method according to claim 21, wherein the immune response is a humoral immune response.

31. A method of enhancing the quality of an immune response to a target antigen in a mammal, which response is conferred by the expression of a parent polynucleotide that encodes a polypeptide corresponding to at least a portion of the target antigen, the method comprising: introducing into the mammal a chimeric construct comprising a synthetic polynucleotide that is operably connected to a regulatory polynucleotide and that is distinguished from the parent polynucleotide by the replacement of a first codon in the parent polynucleotide with a synonymous codon that has a higher immune response preference than the first codon, wherein the first and

synonymous codons are selected according to TABLES 2 or 3, whereby expression of the introduced synthetic polynucleotide confers a stronger or enhanced immune response than the one conferred by expression of the parent polynucleotide under the same conditions.

32. A method of reducing the quality of an immune response to a target antigen in a mammal, which response is conferred by the expression of a parent polynucleotide that encodes a polypeptide corresponding to at least a portion of the target antigen, the method comprising: introducing into the mammal a chimeric construct comprising a synthetic polynucleotide that is operably connected to a regulatory polynucleotide and that is distinguished from the parent polynucleotide by the replacement of a first codon in the parent polynucleotide with a synonymous codon that has a lower immune response preference than the first codon, wherein the first and synonymous codons are selected according to TABLES 5 or 6, whereby expression of the introduced synthetic polynucleotide confers a weaker or reduced immune response than the one conferred by expression of the parent polynucleotide under the same conditions.

33. A method of enhancing the quality of an immune response to a target antigen in a mammal, which response is conferred by the expression of a first polynucleotide that encodes a polypeptide corresponding to at least a portion of the target antigen, the method comprising: co-introducing into the mammal a first nucleic acid construct comprising the first polynucleotide in operable connection with a regulatory polynucleotide; and a second nucleic acid construct comprising a second polynucleotide that is operably connected to a regulatory polynucleotide and that encodes an iso-tRNA corresponding to a codon of the first polynucleotide, wherein the codon has a low or intermediate immune response preference and is selected from the group consisting of Ala^{GCA}, Ala^{GCG}, Ala^{GCC}, Arg^{AGG}, Arg^{CGG}, Asn^{AAT}, Asp^{GAT}, Cys^{TGT}, Glu^{GAG}, Gly^{GGG}, Gly^{GGT}, Gly^{GGC}, Ile^{ATA}, Ile^{ATT}, Leu^{TTG}, Leu^{TTA}, Leu^{CTA}, Leu^{CTT}, Phe^{TTC}, Pro^{CCA}, Pro^{CCG}, Pro^{CCT}, Ser^{AGC}, Ser^{AGT}, Ser^{TCT}, Ser^{TCA}, Ser^{TCC}, Thr^{ACA}, Thr^{ACT}, Tyr^{TAT}, Val^{GTA} and Val^{GTT}.

34. The method according to claim 33, wherein the codon has a 'low' immune response preference, and is selected from the group consisting of Ala^{GCA}, Ala^{GCG}, Arg^{AGG}, Arg^{CGG}, Asn^{AAT}, Asp^{GAT}, Cys^{TGT}, Glu^{GAG}, Gly^{GGG}, Gly^{GGT}, Gly^{GGC}, Ile^{ATA}, Leu^{TTG}, Leu^{TTA}, Phe^{TTC}, Pro^{CCA}, Pro^{CCG}, Ser^{AGC}, Ser^{AGT}, Thr^{ACT}, Tyr^{TAT} and Val^{GTA}.

```
            1         11         21         31         41         51
IgkC1       GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-1     GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-2     GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-3     GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-4     GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkC2       GGTACCGCCGCCACCATGGAGACCGACACCCTCCTGCTGTGGGTGCTGCTGCTCTGGGTG

            61         71         81         91         101        111
IgkC1       CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-1     CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-2     CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-3     CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-4     CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkC2       CCCGGCTCCACCGGCGACGGATCCATGCACGGCGACACCCCCACCCTGCACGAGTACATG

            121        131        141        151        161        171
IgkC1       TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-1     TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-2     TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-3     TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-4     TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkC2       CTGGACCTGCAGCCCGAGACCACCGGCCTGTACGGCTACGGCCAGCTCAACGACAGCAGC

            181        191        201        211        221        231
IgkC1       GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-1     GAGGAGGAGGATGAAATAGATGGTCCAGCGGGACAAGCGGAACCGGACAGAGCGCATTAC
IgkS1-2     GAGGAGGAGGATGAAATAGATGGTCCAGCAGGACAAGCAGAACCGGACAGAGCACATTAC
IgkS1-3     GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCTGAACCGGACAGAGCTCATTAC
IgkS1-4     GAGGAGGAGGATGAAATAGATGGTCCAGCCGGACAAGCCGAACCGGACAGAGCCCATTAC
IgkC2       GAGGAGGAGGACGAGATCGACGGCCCCGCCGGCCAGGCCGAGCCCGACCGCGCCCACTAC

            241        251        261        271        281        291
IgkC1       AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-1     AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-2     AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-3     AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-4     AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkC2       AACATCGTGACCTTCTGCTGCAAGTGCGACAGCACCCTGCGCCTCTGCGTGCAGAGCACC

            301        311        321        331        341        351
IgkC1       CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-1     CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-2     CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-3     CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-4     CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkC2       CACGTGGACATCCGCACCCTGGAGGACCTGCTGATGGGCACCCTGGGCATCGTGTGCCCC

            361        371        381
IgkC1       ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-1     ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-2     ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-3     ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-4     ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkC2       ATCTGCTCCCAGAAGCCCTAAGAATTC
```

FIGURE 1

Substitute Sheet

```
                1         11        21        31        41        51
IgkS1-5         GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-6         GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-7         GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-8         GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-9         GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-10        GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkC1           GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkC2           GGTACCGCCGCCACCATGGAGACCGACACCCTCCTGCTGTGGGTGCTGCTGCTCTGGGTG

                61        71        81        91        101       111
IgkS1-5         CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-6         CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-7         CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-8         CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-9         CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-10        CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkC1           CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkC2           CCCGGCTCCACCGGCGACGGATCCATGCACGGCGACACCCCCACCCTGCACGAGTACATG

                121       131       141       151       161       171
IgkS1-5         TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-6         TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-7         TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-8         TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-9         TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-10        TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkC1           TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkC2           CTGGACCTGCAGCCCGAGACCACCGGCCTGTACGGCTACGGCCAGCTCAACGACAGCAGC

                181       191       201       211       221       231
IgkS1-5         GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGGGCCCATTAC
IgkS1-6         GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-7         GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACCGGGCCCATTAC
IgkS1-8         GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACCGAGCCCATTAC
IgkS1-9         GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACCGTGCCCATTAC
IgkS1-10        GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACCGCGCCCATTAC
IgkC1           GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkC2           GAGGAGGAGGACGAGATCGACGGCCCCGCCGGCCAGGCCGAGCCCGACCGCGCCCACTAC

                241       251       261       271       281       291
IgkS1-5         AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTAGGTTGTGCGTACAAAGCACA
IgkS1-6         AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTAGATTGTGCGTACAAAGCACA
IgkS1-7         AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-8         AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGATTGTGCGTACAAAGCACA
IgkS1-9         AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGTTTGTGCGTACAAAGCACA
IgkS1-10        AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGCTTGTGCGTACAAAGCACA
IgkC1           AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkC2           AACATCGTGACCTTCTGCTGCAAGTGCGACAGCACCCTGCGCCTCTGCGTGCAGAGCACC

                301       311       321       331       341       351
IgkS1-5         CACGTAGACATTAGGACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-6         CACGTAGACATTAGAACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-7         CACGTAGACATTCGGACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-8         CACGTAGACATTCGAACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-9         CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-10        CACGTAGACATTCGCACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkC1           CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkC2           CACGTGGACATCCGCACCCTGGAGGACCTGCTGATGGGCACCCTGGGCATCGTGTGCCCC

                361       371       381
IgkS1-5         ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-6         ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-7         ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-8         ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-9         ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-10        ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkC1           ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkC2           ATCTGCTCCCAGAAGCCCTAAGAATTC
```

FIGURE 2

```
            1         11        21        31        41        51
IgkC1       GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-12    GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-31    GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkC2       GGTACCGCCGCCACCATGGAGACCGACACCCTCCTGCTGTGGGTGCTGCTGCTCTGGGTG

            61        71        81        91        101       111
IgkC1       CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-12    CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-31    CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkC2       CCCGGCTCCACCGGCGACGGATCCATGCACGGCGACACCCCCACCCTGCACGAGTACATG

            121       131       141       151       161       171
IgkC1       TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-12    TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAACGACAGCTCA
IgkS1-31    TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkC2       CTGGACCTGCAGCCCGAGACCACCGGCCTGTACGGCTACGGCCAGCTCAACGACAGCAGC

            181       191       201       211       221       231
IgkC1       GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-12    GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-31    GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkC2       GAGGAGGAGGACGAGATCGACGGCCCCGCCGGCCAGGCCGAGCCCGACCGCGCCCACTAC

            241       251       261       271       281       291
IgkC1       AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-12    AACATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-31    AATATTGTAACCTTTTGTTGCAAATGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkC2       AACATCGTGACCTTCTGCTGCAAGTGCGACAGCACCCTGCGCCTCTGCGTGCAGAGCACC

            301       311       321       331       341       351
IgkC1       CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-12    CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-31    CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkC2       CACGTGGACATCCGCACCCTGGAGGACCTGCTGATGGGCACCCTGGGCATCGTGTGCCCC

            361       371       381
IgkC1       ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-12    ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-31    ATCTGCTCTCAGAAACCCTAAGAATTC
IgkC2       ATCTGCTCCCAGAAGCCCTAAGAATTC
```

FIGURE 3

```
                1         11        21        31        41        51
IgkC1           GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-13        GGTACCGCCGCCACCATGGAGACAGATACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-14        GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkC2           GGTACCGCCGCCACCATGGAGACCGACACCCTCCTGCTGTGGGTGCTGCTGCTCTGGGTG

                61        71        81        91        101       111
IgkC1           CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-13        CCAGGTTCCACTGGTGATGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-14        CCAGGTTCCACTGGTGACGGATCCATGCATGGAGACACACCTACATTGCATGAATATATG
IgkC2           CCCGGCTCCACCGGCGACGGATCCATGCACGGCGACACCCCCACCCTGCACGAGTACATG

                121       131       141       151       161       171
IgkC1           TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-13        TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGATAGCTCA
IgkS1-14        TTAGACTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkC2           CTGGACCTGCAGCCCGAGACCACCGGCCTGTACGGCTACGGCCAGCTCAACGACAGCAGC

                181       191       201       211       221       231
IgkC1           GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-13        GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGATAGAGCCCATTAC
IgkS1-14        GAGGAGGAGGACGAAATAGACGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkC2           GAGGAGGAGGACGAGATCGACGGCCCCGCCGGCCAGGCCGAGCCCGACCGCGCCCACTAC

                241       251       261       271       281       291
IgkC1           AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-13        AATATTGTAACCTTTTGTTGCAAGTGTGATTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-14        AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkC2           AACATCGTGACCTTCTGCTGCAAGTGCGACAGCACCCTGCGCCTCTGCGTGCAGAGCACC

                301       311       321       331       341       351
IgkC1           CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-13        CACGTAGATATTCGTACTTTGGAAGATCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-14        CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkC2           CACGTGGACATCCGCACCCTGGAGGACCTGCTGATGGGCACCCTGGGCATCGTGTGCCCC

                361       371       381
IgkC1           ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-13        ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-14        ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkC2           ATCTGCTCCCAGAAGCCCTAAGAATTC
```

FIGURE 4

```
                1         11         21         31         41         51
IgkC1           GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-15        GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-16        GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkC2           GGTACCGCCGCCACCATGGAGACCGACACCCTCCTGCTGTGGGTGCTGCTGCTCTGGGTG

                61         71         81         91         101        111
IgkC1           CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-15        CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-16        CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkC2           CCCGGCTCCACCGGCGACGGATCCATGCACGGCGACACCCCCACCCTGCACGAGTACATG

                121        131        141        151        161        171
IgkC1           TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-15        TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-16        TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkC2           CTGGACCTGCAGCCCGAGACCACCGGCCTGTACGGCTACGGCCAGCTCAACGACAGCAGC

                181        191        201        211        221        231
IgkC1           GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-15        GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-16        GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkC2           GAGGAGGAGGACGAGATCGACGGCCCCGCCGGCCAGGCCGAGCCCGACCGCGCCCACTAC

                241        251        261        271        281        291
IgkC1           AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-15        AATATTGTAACCTTTTGTTGTAAGTGTGACTCTACGCTTCGGTTGTGTGTACAAAGCACA
IgkS1-16        AATATTGTAACCTTTTGCTGCAAGTGCGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkC2           AACATCGTGACCTTCTGCTGCAAGTGCGACAGCACCCTGCGCCTCTGCGTGCAGAGCACC

                301        311        321        331        341        351
IgkC1           CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-15        CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGTCCC
IgkS1-16        CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkC2           CACGTGGACATCCGCACCCTGGAGGACCTGCTGATGGGCACCCTGGGCATCGTGTGCCCC

                361        371        381
IgkC1           ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-15        ATCTGTTCTCAGAAGCCCTAAGAATTC
IgkS1-16        ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkC2           ATCTGCTCCCAGAAGCCCTAAGAATTC
```

FIGURE 5

```
                1         11        21        31        41        51
IgkS1-17        GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-18        GGTACCGCCGCCACCATGGAAACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkC2           GGTACCGCCGCCACCATGGAGACCGACACCCTCCTGCTGTGGGTGCTGCTGCTCTGGGTG
IgkC1           GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT

                61        71        81        91        101       111
IgkS1-17        CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAGTATATG
IgkS1-18        CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkC2           CCCGGCTCCACCGGCGACGGATCCATGCACGGCGACACCCCCACCCTGCACGAGTACATG
IgkC1           CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG

                121       131       141       151       161       171
IgkS1-17        TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-18        TTAGATTTGCAACCAGAAACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkC2           CTGGACCTGCAGCCCGAGACCACCGGCCTGTACGGCTACGGCCAGCTCAACGACAGCAGC
IgkC1           TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA

                181       191       201       211       221       231
IgkS1-17        GAGGAGGAGGATGAGATAGATGGTCCAGCTGGACAAGCAGAGCCGGACAGAGCCCATTAC
IgkS1-18        GAAGAAGAAGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkC2           GAGGAGGAGGACGAGATCGACGGCCCCGCCGGCCAGGCCGAGCCCGACCGCGCCCACTAC
IgkC1           GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC

                241       251       261       271       281       291
IgkS1-17        AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-18        AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkC2           AACATCGTGACCTTCTGCTGCAAGTGCGACAGCACCCTGCGCCTCTGCGTGCAGAGCACC
IgkC1           AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA

                301       311       321       331       341       351
IgkS1-17        CACGTAGACATTCGTACTTTGGAGGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-18        CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkC2           CACGTGGACATCCGCACCCTGGAGGACCTGCTGATGGGCACCCTGGGCATCGTGTGCCCC
IgkC1           CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC

                361       371       381
IgkS1-17        ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-18        ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkC2           ATCTGCTCCCAGAAGCCCTAAGAATTC
IgkC1           ATCTGCTCTCAGAAGCCCTAAGAATTC
```

FIGURE 6

```
                1         11        21        31        41        51
IgkC1           GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-19        GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-20        GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkC2           GGTACCGCCGCCACCATGGAGACCGACACCCTCCTGCTGTGGGTGCTGCTGCTCTGGGTG

                61        71        81        91        101       111
IgkC1           CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-19        CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-20        CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkC2           CCCGGCTCCACCGGCGACGGATCCATGCACGGCGACACCCCCACCCTGCACGAGTACATG

                121       131       141       151       161       171
IgkC1           TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-19        TTAGATTTGCAGCCAGAGACAACTGGTCTCTACGGTTATGGGCAGTTAAATGACAGCTCA
IgkS1-20        TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkC2           CTGGACCTGCAGCCCGAGACCACCGGCCTGTACGGCTACGGCCAGCTCAACGACAGCAGC

                181       191       201       211       221       231
IgkC1           GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-19        GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAGGCAGAACCGGACAGAGCCCATTAC
IgkS1-20        GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkC2           GAGGAGGAGGACGAGATCGACGGCCCCGCCGGCCAGGCCGAGCCCGACCGCGCCCACTAC

                241       251       261       271       281       291
IgkC1           AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-19        AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAGAGCACA
IgkS1-20        AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkC2           AACATCGTGACCTTCTGCTGCAAGTGCGACAGCACCCTGCGCCTCTGCGTGCAGAGCACC

                301       311       321       331       341       351
IgkC1           CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-19        CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-20        CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkC2           CACGTGGACATCCGCACCCTGGAGGACCTGCTGATGGGCACCCTGGGCATCGTGTGCCCC

                361       371       381
IgkC1           ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-19        ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-20        ATCTGCTCTCAAAAGCCCTAAGAATTC
IgkC2           ATCTGCTCCCAGAAGCCCTAAGAATTC
```

FIGURE 7

```
                1         11        21        31        41        51
IgkC1           GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-21        GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-22        GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-23        GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-24        GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkC2           GGTACCGCCGCCACCATGGAGACCGACACCCTCCTGCTGTGGGTGCTGCTGCTCTGGGTG

                61        71        81        91        101       111
IgkC1           CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-21        CCAGGGTCCACTGGGGACGGATCCATGCATGGGGATACACCTACATTGCATGAATATATG
IgkS1-22        CCAGGATCCACTGGAGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-23        CCAGGTTCCACTGGTGACGGATCCATGCATGGTGATACACCTACATTGCATGAATATATG
IgkS1-24        CCAGGCTCCACTGGCGACGGATCCATGCATGGCGATACACCTACATTGCATGAATATATG
IgkC2           CCCGGCTCCACCGGCGACGGATCCATGCACGGCGACACCCCCACCCTGCACGAGTACATG

                121       131       141       151       161       171
IgkC1           TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-21        TTAGATTTGCAACCAGAGACAACTGGGCTCTACGGGTATGGGCAATTAAATGACAGCTCA
IgkS1-22        TTAGATTTGCAACCAGAGACAACTGGACTCTACGGATATGGACAATTAAATGACAGCTCA
IgkS1-23        TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGTCAATTAAATGACAGCTCA
IgkS1-24        TTAGATTTGCAACCAGAGACAACTGGCCTCTACGGCTATGGCCAATTAAATGACAGCTCA
IgkC2           CTGGACCTGCAGCCCGAGACCACCGGCCTGTACGGCTACGGCCAGCTCAACGACAGCAGC

                181       191       201       211       221       231
IgkC1           GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-21        GAGGAGGAGGATGAAATAGATGGGCCAGCTGGGCAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-22        GAGGAGGAGGATGAAATAGATGGACCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-23        GAGGAGGAGGATGAAATAGATGGTCCAGCTGGTCAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-24        GAGGAGGAGGATGAAATAGATGGCCCAGCTGGCCAAGCAGAACCGGACAGAGCCCATTAC
IgkC2           GAGGAGGAGGACGAGATCGACGGCCCCGCCGGCCAGGCCGAGCCCGACCGCGCCCACTAC

                241       251       261       271       281       291
IgkC1           AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-21        AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-22        AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-23        AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-24        AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkC2           AACATCGTGACCTTCTGCTGCAAGTGCGACAGCACCCTGCGCCTCTGCGTGCAGAGCACC

                301       311       321       331       341       351
IgkC1           CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-21        CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGGACACTAGGGATTGTGTGCCCC
IgkS1-22        CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGAACACTAGGAATTGTGTGCCCC
IgkS1-23        CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGTACACTAGGTATTGTGTGCCCC
IgkS1-24        CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGCATTGTGTGCCCC
IgkC2           CACGTGGACATCCGCACCCTGGAGGACCTGCTGATGGGCACCCTGGGCATCGTGTGCCCC

                361       371       381
IgkC1           ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-21        ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-22        ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-23        ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-24        ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkC2           ATCTGCTCCCAGAAGCCCTAAGAATTC
```

FIGURE 8

```
                1         11        21        31        41        51
IgkC1           GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-25        GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-26        GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkC2           GGTACCGCCGCCACCATGGAGACCGACACCCTCCTGCTGTGGGTGCTGCTGCTCTGGGTG

                61        71        81        91        101       111
IgkC1           CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-25        CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-26        CCAGGTTCCACTGGTGACGGATCCATGCACGGAGATACACCTACATTGCACGAATATATG
IgkC2           CCCGGCTCCACCGGCGACGGATCCATGCACGGCGACACCCCCACCCTGCACGAGTACATG

                121       131       141       151       161       171
IgkC1           TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-25        TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-26        TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkC2           CTGGACCTGCAGCCCGAGACCACCGGCCTGTACGGCTACGGCCAGCTCAACGACAGCAGC

                181       191       201       211       221       231
IgkC1           GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-25        GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-26        GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCACTAC
IgkC2           GAGGAGGAGGACGAGATCGACGGCCCCGCCGGCCAGGCCGAGCCCGACCGCGCCCACTAC

                241       251       261       271       281       291
IgkC1           AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-25        AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-26        AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkC2           AACATCGTGACCTTCTGCTGCAAGTGCGACAGCACCCTGCGCCTCTGCGTGCAGAGCACC

                301       311       321       331       341       351
IgkC1           CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-25        CATGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-26        CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkC2           CACGTGGACATCCGCACCCTGGAGGACCTGCTGATGGGCACCCTGGGCATCGTGTGCCCC

                361       371       381
IgkC1           ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-25        ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-26        ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkC2           ATCTGCTCCCAGAAGCCCTAAGAATTC
```

FIGURE 9

```
                1         11        21        31        41        51
IgkC1           GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-27        GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-28        GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-29        GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkC2           GGTACCGCCGCCACCATGGAGACCGACACCCTCCTGCTGTGGGTGCTGCTGCTCTGGGTG

                61        71        81        91        101       111
IgkC1           CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-27        CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-28        CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-29        CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkC2           CCCGGCTCCACCGGCGACGGATCCATGCACGGCGACACCCCCACCCTGCACGAGTACATG

                121       131       141       151       161       171
IgkC1           TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-27        TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-28        TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-29        TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkC2           CTGGACCTGCAGCCCGAGACCACCGGCCTGTACGGCTACGGCCAGCTCAACGACAGCAGC

                181       191       201       211       221       231
IgkC1           GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-27        GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-28        GAGGAGGAGGATGAAATTGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-29        GAGGAGGAGGATGAAATCGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkC2           GAGGAGGAGGACGAGATCGACGGCCCCGCCGGCCAGGCCGAGCCCGACCGCGCCCACTAC

                241       251       261       271       281       291
IgkC1           AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-27        AATATAGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-28        AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-29        AATATCGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkC2           AACATCGTGACCTTCTGCTGCAAGTGCGACAGCACCCTGCGCCTCTGCGTGCAGAGCACC

                301       311       321       331       341       351
IgkC1           CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-27        CACGTAGACATACGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATAGTGTGCCCC
IgkS1-28        CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-29        CACGTAGACATCCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATCGTGTGCCCC
IgkC2           CACGTGGACATCCGCACCCTGGAGGACCTGCTGATGGGCACCCTGGGCATCGTGTGCCCC

                361       371       381
IgkC1           ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-27        ATATGCTCTCAGAAGCCCTAAGAATTC
IgkS1-28        ATTTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-29        ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkC2           ATCTGCTCCCAGAAGCCCTAAGAATTC
```

FIGURE 10

```
                1         11        21        31        41        51
IgkS1-50        GGTACCGCCGCCACCATGGAAACTGACACTCTGCTGCTGTGGGTACTGCTGCTGTGGGTT
IgkS1-51        GGTACCGCCGCCACCATGGAAACTGACACTCTACTACTATGGGTACTACTACTATGGGTT
IgkS1-52        GGTACCGCCGCCACCATGGAAACTGACACTCTTCTTCTTTGGGTACTTCTTCTTTGGGTT
IgkS1-53        GGTACCGCCGCCACCATGGAAACTGACACTCTCCTCCTCTGGGTACTCCTCCTCTGGGTT
IgkS1-54        GGTACCGCCGCCACCATGGAAACTGACACTTTGTTGTTGTGGGTATTGTTGTTGTGGGTT
IgkS1-55        GGTACCGCCGCCACCATGGAAACTGACACTTTATTATTATGGGTATTATTATTATGGGTT
IgkC3           GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkC4           GGTACCGCCGCCACCATGGAGACCGACACCCTCCTGCTGTGGGTGCTGCTGCTCTGGGTG

                61        71        81        91        101       111
IgkS1-50        CCAGGATCGACTGGAGACGGATCCATGCATGGAGACACTCCAACTCTGCATGAATATATG
IgkS1-51        CCAGGATCGACTGGAGACGGATCCATGCATGGAGACACTCCAACTCTACATGAATATATG
IgkS1-52        CCAGGATCGACTGGAGACGGATCCATGCATGGAGACACTCCAACTCTTCATGAATATATG
IgkS1-53        CCAGGATCGACTGGAGACGGATCCATGCATGGAGACACTCCAACTCTCCATGAATATATG
IgkS1-54        CCAGGATCGACTGGAGACGGATCCATGCATGGAGACACTCCAACTTTGCATGAATATATG
IgkS1-55        CCAGGATCGACTGGAGACGGATCCATGCATGGAGACACTCCAACTTTACATGAATATATG
IgkC3           CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkC4           CCCGGCTCCACCGGCGACGGATCCATGCACGGCGACACCCCCACCCTGCACGAGTACATG

                121       131       141       151       161       171
IgkS1-50        CTGGACCTGCAACCGGAAACTACTGACCTGTACTGCTATGAACAACTGAATGACAGCTCG
IgkS1-51        CTAGACCTACAACCGGAAACTACTGACCTATACTGCTATGAACAACTAAATGACAGCTCG
IgkS1-52        CTTGACCTTCAACCGGAAACTACTGACCTTTACTGCTATGAACAACTTAATGACAGCTCG
IgkS1-53        CTCGACCTCCAACCGGAAACTACTGACCTCTACTGCTATGAACAACTCAATGACAGCTCG
IgkS1-54        TTGGACTTGCAACCGGAAACTACTGACTTGTACTGCTATGAACAATTGAATGACAGCTCG
IgkS1-55        TTAGACTTACAACCGGAAACTACTGACTTATACTGCTATGAACAATTAAATGACAGCTCG
IgkC3           TTAGATTTGCAACCAGAGACAACTGATCTCTACTGTTATGAGCAATTAAATGACAGCTCA
IgkC4           CTGGACCTGCAGCCCGAGACCACCGACCTGTACTGCTACGAGCAGCTCAACGACAGCAGC

                181       191       201       211       221       231
IgkS1-50        GAAGAAGAAGACGAAATAGACGGACCTGCAGGACAAGCAGAACCAGACCGCGCACATTAC
IgkS1-51        GAAGAAGAAGACGAAATAGACGGACCTGCAGGACAAGCAGAACCAGACCGCGCACATTAC
IgkS1-52        GAAGAAGAAGACGAAATAGACGGACCTGCAGGACAAGCAGAACCAGACCGCGCACATTAC
IgkS1-53        GAAGAAGAAGACGAAATAGACGGACCTGCAGGACAAGCAGAACCAGACCGCGCACATTAC
IgkS1-54        GAAGAAGAAGACGAAATAGACGGACCTGCAGGACAAGCAGAACCAGACCGCGCACATTAC
IgkS1-55        GAAGAAGAAGACGAAATAGACGGACCTGCAGGACAAGCAGAACCAGACCGCGCACATTAC
IgkC3           GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkC4           GAGGAGGAGGACGAGATCGACGGCCCCGCCGGCCAGGCCGAGCCCGACCGCGCCCACTAC

                241       251       261       271       281       291

IgkS1-50        AATATTGTAACTTTTTGCTGCAAGTGCGACAGTACTCTGCGCCTGTGCGTACAAAGCACT
IgkS1-51        AATATTGTAACTTTTTGCTGCAAGTGCGACAGTACTCTACGCCTATGCGTACAAAGCACT
IgkS1-52        AATATTGTAACTTTTTGCTGCAAGTGCGACAGTACTCTTCGCCTTTGCGTACAAAGCACT
IgkS1-53        AATATTGTAACTTTTTGCTGCAAGTGCGACAGTACTCTCCGCCTCTGCGTACAAAGCACT
IgkS1-54        AATATTGTAACTTTTTGCTGCAAGTGCGACAGTACTTTGCGCTTGTGCGTACAAAGCACT
IgkS1-55        AATATTGTAACTTTTTGCTGCAAGTGCGACAGTACTTTACGCTTATGCGTACAAAGCACT
IgkC3           AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkC4           AACATCGTGACCTTCTGCTGCAAGTGCGACAGCACCCTGCGCCTCTGCGTGCAGAGCACC

                301       311       321       331       341       351
IgkS1-50        CATGTAGACATTCGCACTCTGGAAGACCTGCTGATGGGAACTCTGGGAATTGTTTGCCCG
IgkS1-51        CATGTAGACATTCGCACTCTAGAAGACCTACTAATGGGAACTCTAGGAATTGTTTGCCCG
IgkS1-52        CATGTAGACATTCGCACTCTTGAAGACCTTCTTATGGGAACTCTTGGAATTGTTTGCCCG
IgkS1-53        CATGTAGACATTCGCACTCTCGAAGACCTCCTCATGGGAACTCTCGGAATTGTTTGCCCG
IgkS1-54        CATGTAGACATTCGCACTTTGGAAGACTTGTTGATGGGAACTTTGGGAATTGTTTGCCCG
IgkS1-55        CATGTAGACATTCGCACTTTAGAAGACTTATTAATGGGAACTTTAGGAATTGTTTGCCCG
IgkC3           CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkC4           CACGTGGACATCCGCACCCTGGAGGACCTGCTGATGGGCACCCTGGGCATCGTGTGCCCC

                361       371       381
IgkS1-51        ATCTGCTCGCAAAAGCCTTAAGAATTC
IgkS1-53        ATCTGCTCGCAAAAGCCTTAAGAATTC
IgkS1-50        ATCTGCTCGCAAAAGCCTTAAGAATTC
IgkS1-52        ATCTGCTCGCAAAAGCCTTAAGAATTC
IgkS1-55        ATCTGCTCGCAAAAGCCTTAAGAATTC
IgkS1-54        ATCTGCTCGCAAAAGCCTTAAGAATTC
IgkC3           ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkC4           ATCTGCTCCCAGAAGCCCTAAGAATTC
```

FIGURE 11

```
                1         11        21        31        41        51
IgkS1-32        GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-33        GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkC1           GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkC2           GGTACCGCCGCCACCATGGAGACCGACACCCTCCTGCTGTGGGTGCTGCTGCTCTGGGTG

                61        71        81        91        101       111
IgkS1-32        CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-33        CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkC1           CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkC2           CCCGGCTCCACCGGCGACGGATCCATGCACGGCGACACCCCCACCCTGCACGAGTACATG

                121       131       141       151       161       171
IgkS1-32        TTAGATTTTCAACCAGAGACAACTGGTTTTTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-33        TTAGATTTCCAACCAGAGACAACTGGTTTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkC1           TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkC2           CTGGACCTGCAGCCCGAGACCACCGGCCTGTACGGCTACGGCCAGCTCAACGACAGCAGC

                181       191       201       211       221       231
IgkS1-32        GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-33        GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkC1           GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkC2           GAGGAGGAGGACGAGATCGACGGCCCCGCCGGCCAGGCCGAGCCCGACCGCGCCCACTAC

                241       251       261       271       281       291
IgkS1-32        AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-33        AATATTGTAACCTTCTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkC1           AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkC2           AACATCGTGACCTTCTGCTGCAAGTGCGACAGCACCCTGCGCCTCTGCGTGCAGAGCACC

                301       311       321       331       341       351
IgkS1-32        CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-33        CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkC1           CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkC2           CACGTGGACATCCGCACCCTGGAGGACCTGCTGATGGGCACCCTGGGCATCGTGTGCCCC

                361       371       381
IgkS1-32        ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-33        ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkC1           ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkC2           ATCTGCTCCCAGAAGCCCTAAGAATTC
```

FIGURE 12

```
                1          11         21         31         41         51
IgkS1-56        GGTACCGCCGCCACCATGGAAACTGACACTCTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-57        GGTACCGCCGCCACCATGGAAACTGACACTCTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-58        GGTACCGCCGCCACCATGGAAACTGACACTCTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-59        GGTACCGCCGCCACCATGGAAACTGACACTCTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkC3           GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkC4           GGTACCGCCGCCACCATGGAGACCGACACCCTCCTGCTGTGGGTGCTGCTGCTCTGGGTG

                61         71         81         91         101        111
IgkS1-56        CCGGGATCGACTGGAGACGGATCCATGCATGGAGACACTCCGACTTTGCATGAATATATG
IgkS1-57        CCAGGATCGACTGGAGACGGATCCATGCATGGAGACACTCCAACTTTGCATGAATATATG
IgkS1-58        CCTGGATCGACTGGAGACGGATCCATGCATGGAGACACTCCTACTTTGCATGAATATATG
IgkS1-59        CCCGGATCGACTGGAGACGGATCCATGCATGGAGACACTCCCACTTTGCATGAATATATG
IgkC3           CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkC4           CCCGGCTCCACCGGCGACGGATCCATGCACGGCGACACCCCCACCCTGCACGAGTACATG

                121        131        141        151        161        171
IgkS1-56        CTCGACTTGCAACCGGAAACTACTGACCTCTACTGCTATGAACAATTGAATGACAGCTCG
IgkS1-57        CTCGACTTGCAACCAGAAACTACTGACCTCTACTGCTATGAACAATTGAATGACAGCTCG
IgkS1-58        CTCGACTTGCAACCTGAAACTACTGACCTCTACTGCTATGAACAATTGAATGACAGCTCG
IgkS1-59        CTCGACTTGCAACCCGAAACTACTGACCTCTACTGCTATGAACAATTGAATGACAGCTCG
IgkC3           TTAGATTTGCAACCAGAGACAACTGATCTCTACTGTTATGAGCAATTAAATGACAGCTCA
IgkC4           CTGGACCTGCAGCCCGAGACCACCGACCTGTACTGCTACGAGCAGCTCAACGACAGCAGC

                181        191        201        211        221        231
IgkS1-56        GAAGAAGAAGACGAAATAGACGGACCGGCAGGACAAGCAGAACCGGACCGCGCACATTAC
IgkS1-57        GAAGAAGAAGACGAAATAGACGGACCAGCAGGACAAGCAGAACCAGACCGCGCACATTAC
IgkS1-58        GAAGAAGAAGACGAAATAGACGGACCTGCAGGACAAGCAGAACCTGACCGCGCACATTAC
IgkS1-59        GAAGAAGAAGACGAAATAGACGGACCCGCAGGACAAGCAGAACCCGACCGCGCACATTAC
IgkC3           GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkC4           GAGGAGGAGGACGAGATCGACGGCCCCGCCGGCCAGGCCGAGCCCGACCGCGCCCACTAC

                241        251        261        271        281        291
IgkS1-56        AATATTGTAACTTTTTGCTGCAAGTGCGACAGTACTCTCCGCTTGTGCGTACAAAGCACT
IgkS1-57        AATATTGTAACTTTTTGCTGCAAGTGCGACAGTACTCTCCGCTTGTGCGTACAAAGCACT
IgkS1-58        AATATTGTAACTTTTTGCTGCAAGTGCGACAGTACTCTCCGCTTGTGCGTACAAAGCACT
IgkS1-59        AATATTGTAACTTTTTGCTGCAAGTGCGACAGTACTCTCCGCTTGTGCGTACAAAGCACT
IgkC3           AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkC4           AACATCGTGACCTTCTGCTGCAAGTGCGACAGCACCCTGCGCCTCTGCGTGCAGAGCACC

                301        311        321        331        341        351
IgkS1-56        CATGTAGACATTCGCACTTTGGAAGACCTCCTCATGGGAACTTTGGGAATTGTTTGCCCG
IgkS1-57        CATGTAGACATTCGCACTTTGGAAGACCTCCTCATGGGAACTTTGGGAATTGTTTGCCCA
IgkS1-58        CATGTAGACATTCGCACTTTGGAAGACCTCCTCATGGGAACTTTGGGAATTGTTTGCCCT
IgkS1-59        CATGTAGACATTCGCACTTTGGAAGACCTCCTCATGGGAACTTTGGGAATTGTTTGCCCC
IgkC3           CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkC4           CACGTGGACATCCGCACCCTGGAGGACCTGCTGATGGGCACCCTGGGCATCGTGTGCCCC

                361        371        381
IgkS1-56        ATCTGCTCGCAAAAGCCGTAAGAATTC
IgkS1-57        ATCTGCTCGCAAAAGCCATAAGAATTC
IgkS1-58        ATCTGCTCGCAAAAGCCTTAAGAATTC
IgkS1-59        ATCTGCTCGCAAAAGCCCTAAGAATTC
IgkC3           ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkC4           ATCTGCTCCCAGAAGCCCTAAGAATTC
```

FIGURE 13

```
                1         11         21         31         41         51
IgkS1-34        GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-35        GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-36        GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-37        GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-38        GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-39        GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkC1           GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkC2           GGTACCGCCGCCACCATGGAGACCGACACCCTCCTGCTGTGGGTGCTGCTGCTCTGGGTG

                61         71         81         91         101        111
IgkS1-34        CCAGGTAGTACTGGTGACGGAAGTATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-35        CCAGGTAGCACTGGTGACGGAAGCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-36        CCAGGTTCGACTGGTGACGGATCGATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-37        CCAGGTTCAACTGGTGACGGATCAATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-38        CCAGGTTCTACTGGTGACGGATCTATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-39        CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkC1           CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkC2           CCCGGCTCCACCGGCGACGGATCCATGCACGGCGACACCCCCACCCTGCACGAGTACATG

                121        131        141        151        161        171
IgkS1-34        TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGTAGT
IgkS1-35        TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCAGC
IgkS1-36        TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACTCGTCG
IgkS1-37        TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACTCATCA
IgkS1-38        TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACTCTTCT
IgkS1-39        TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACTCCTCC
IgkC1           TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkC2           CTGGACCTGCAGCCCGAGACCACCGGCCTGTACGGCTACGGCCAGCTCAACGACAGCAGC

                181        191        201        211        221        231
IgkS1-34        GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-35        GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-36        GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-37        GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-38        GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-39        GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkC1           GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkC2           GAGGAGGAGGACGAGATCGACGGCCCCGCCGGCCAGGCCGAGCCCGACCGCGCCCACTAC

                241        251        261        271        281        291
IgkS1-34        AATATTGTAACCTTTTGTTGCAAGTGTGACAGTACGCTTCGGTTGTGCGTACAAAGTACA
IgkS1-35        AATATTGTAACCTTTTGTTGCAAGTGTGACAGCACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-36        AATATTGTAACCTTTTGTTGCAAGTGTGACTCGACGCTTCGGTTGTGCGTACAATCGACA
IgkS1-37        AATATTGTAACCTTTTGTTGCAAGTGTGACTCAACGCTTCGGTTGTGCGTACAATCAACA
IgkS1-38        AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAATCTACA
IgkS1-39        AATATTGTAACCTTTTGTTGCAAGTGTGACTCCACGCTTCGGTTGTGCGTACAATCCACA
IgkC1           AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkC2           AACATCGTGACCTTCTGCTGCAAGTGCGACAGCACCCTGCGCCTCTGCGTGCAGAGCACC

                301        311        321        331        341        351
IgkS1-34        CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-35        CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-36        CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-37        CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-38        CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-39        CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkC1           CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkC2           CACGTGGACATCCGCACCCTGGAGGACCTGCTGATGGGCACCCTGGGCATCGTGTGCCCC

                361        371        381
IgkS1-34        ATCTGCAGTCAGAAGCCCTAAGAATTC
IgkS1-35        ATCTGCAGCCAGAAGCCCTAAGAATTC
IgkS1-36        ATCTGCTCGCAGAAGCCCTAAGAATTC
IgkS1-37        ATCTGCTCACAGAAGCCCTAAGAATTC
IgkS1-38        ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-39        ATCTGCTCCCAGAAGCCCTAAGAATTC
IgkC1           ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkC2           ATCTGCTCCCAGAAGCCCTAAGAATTC
```

FIGURE 14

```
                1          11         21         31         41         51
IgkC1           GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-40        GGTACCGCCGCCACCATGGAGACGGACACGCTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-41        GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-42        GGTACCGCCGCCACCATGGAGACTGACACTCTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-43        GGTACCGCCGCCACCATGGAGACCGACACCCTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkC2           GGTACCGCCGCCACCATGGAGACCGACACCCTCCTGCTGTGGGTGCTGCTGCTCTGGGTG

                61         71         81         91         101        111
IgkC1           CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-40        CCAGGTTCCACGGGTGACGGATCCATGCATGGAGATACGCCTACGTTGCATGAATATATG
IgkS1-41        CCAGGTTCCACAGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-42        CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACTCCTACTTTGCATGAATATATG
IgkS1-43        CCAGGTTCCACCGGTGACGGATCCATGCATGGAGATACCCCTACCTTGCATGAATATATG
IgkC2           CCCGGCTCCACCGGCGACGGATCCATGCACGGCGACACCCCCACCCTGCACGAGTACATG

                121        131        141        151        161        171
IgkC1           TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-40        TTAGATTTGCAACCAGAGACGACGGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-41        TTAGATTTGCAACCAGAGACAACAGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-42        TTAGATTTGCAACCAGAGACTACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-43        TTAGATTTGCAACCAGAGACCACCGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkC2           CTGGACCTGCAGCCCGAGACCACCGGCCTGTACGGCTACGGCCAGCTCAACGACAGCAGC

                181        191        201        211        221        231
IgkC1           GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-40        GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-41        GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-42        GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-43        GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkC2           GAGGAGGAGGACGAGATCGACGGCCCCGCCGGCCAGGCCGAGCCCGACCGCGCCCACTAC

                241        251        261        271        281        291
IgkC1           AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-40        AATATTGTAACGTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACG
IgkS1-41        AATATTGTAACATTTTGTTGCAAGTGTGACTCTACACTTCGGTTGTGCGTACAAAGCACA
IgkS1-42        AATATTGTAACTTTTTGTTGCAAGTGTGACTCTACTCTTCGGTTGTGCGTACAAAGCACT
IgkS1-43        AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACCCTTCGGTTGTGCGTACAAAGCACC
IgkC2           AACATCGTGACCTTCTGCTGCAAGTGCGACAGCACCCTGCGCCTCTGCGTGCAGAGCACC

                301        311        321        331        341        351
IgkC1           CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-40        CACGTAGACATTCGTACGTTGGAAGACCTGTTAATGGGCACGCTAGGAATTGTGTGCCCC
IgkS1-41        CACGTAGACATTCGTACATTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-42        CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACTCTAGGAATTGTGTGCCCC
IgkS1-43        CACGTAGACATTCGTACCTTGGAAGACCTGTTAATGGGCACCCTAGGAATTGTGTGCCCC
IgkC2           CACGTGGACATCCGCACCCTGGAGGACCTGCTGATGGGCACCCTGGGCATCGTGTGCCCC

                361        371        381
IgkC1           ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-40        ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-41        ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-42        ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-43        ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkC2           ATCTGCTCCCAGAAGCCCTAAGAATTC
```

FIGURE 15

```
            1         11        21        31        41        51
IgkC1       GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-44    GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-45    GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkC2       GGTACCGCCGCCACCATGGAGACCGACACCCTCCTGCTGTGGGTGCTGCTGCTCTGGGTG

            61        71        81        91        101       111
IgkC1       CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-44    CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-45    CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATACATG
IgkC2       CCCGGCTCCACCGGCGACGGATCCATGCACGGCGACACCCCCACCCTGCACGAGTACATG

            121       131       141       151       161       171
IgkC1       TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-44    TTAGATTTGCAACCAGAGACAACTGGTCTCTATGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-45    TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTACGGGCAATTAAATGACAGCTCA
IgkC2       CTGGACCTGCAGCCCGAGACCACCGGCCTGTACGGCTACGGCCAGCTCAACGACAGCAGC

            181       191       201       211       221       231
IgkC1       GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-44    GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAT
IgkS1-45    GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkC2       GAGGAGGAGGACGAGATCGACGGCCCCGCCGGCCAGGCCGAGCCCGACCGCGCCCACTAC

            241       251       261       271       281       291
IgkC1       AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-44    AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-45    AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkC2       AACATCGTGACCTTCTGCTGCAAGTGCGACAGCACCCTGCGCCTCTGCGTGCAGAGCACC

            301       311       321       331       341       351
IgkC1       CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-44    CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-45    CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkC2       CACGTGGACATCCGCACCCTGGAGGACCTGCTGATGGGCACCCTGGGCATCGTGTGCCCC

            361       371       381
IgkC1       ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-44    ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-45    ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkC2       ATCTGCTCCCAGAAGCCCTAAGAATTC
```

FIGURE 16

```
            1         11        21        31        41        51
IgkC1       GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTT
IgkS1-46    GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTGCTGCTGCTCTGGGTG
IgkS1-47    GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTACTGCTGCTCTGGGTA
IgkS1-48    GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTTCTGCTGCTCTGGGTT
IgkS1-49    GGTACCGCCGCCACCATGGAGACAGACACACTCCTGCTATGGGTCCTGCTGCTCTGGGTC
IgkC2       GGTACCGCCGCCACCATGGAGACCGACACCCTCCTGCTGTGGGTGCTGCTGCTCTGGGTG

            61        71        81        91        101       111
IgkC1       CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-46    CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-47    CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-48    CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkS1-49    CCAGGTTCCACTGGTGACGGATCCATGCATGGAGATACACCTACATTGCATGAATATATG
IgkC2       CCCGGCTCCACCGGCGACGGATCCATGCACGGCGACACCCCCACCCTGCACGAGTACATG

            121       131       141       151       161       171
IgkC1       TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-46    TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-47    TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-48    TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkS1-49    TTAGATTTGCAACCAGAGACAACTGGTCTCTACGGTTATGGGCAATTAAATGACAGCTCA
IgkC2       CTGGACCTGCAGCCCGAGACCACCGGCCTGTACGGCTACGGCCAGCTCAACGACAGCAGC

            181       191       201       211       221       231
IgkC1       GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-46    GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-47    GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-48    GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkS1-49    GAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTAC
IgkC2       GAGGAGGAGGACGAGATCGACGGCCCCGCCGGCCAGGCCGAGCCCGACCGCGCCCACTAC

            241       251       261       271       281       291
IgkC1       AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-46    AATATTGTGACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTGCAAAGCACA
IgkS1-47    AATATTGTAACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTACAAAGCACA
IgkS1-48    AATATTGTTACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTTCAAAGCACA
IgkS1-49    AATATTGTCACCTTTTGTTGCAAGTGTGACTCTACGCTTCGGTTGTGCGTCCAAAGCACA
IgkC2       AACATCGTGACCTTCTGCTGCAAGTGCGACAGCACCCTGCGCCTCTGCGTGCAGAGCACC

            301       311       321       331       341       351
IgkC1       CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-46    CACGTGGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCC
IgkS1-47    CACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTATGCCCC
IgkS1-48    CACGTTGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTTTGCCCC
IgkS1-49    CACGTCGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTCTGCCCC
IgkC2       CACGTGGACATCCGCACCCTGGAGGACCTGCTGATGGGCACCCTGGGCATCGTGTGCCCC

            361       371       381
IgkC1       ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-46    ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-47    ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-48    ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkS1-49    ATCTGCTCTCAGAAGCCCTAAGAATTC
IgkC2       ATCTGCTCCCAGAAGCCCTAAGAATTC
```

FIGURE 17

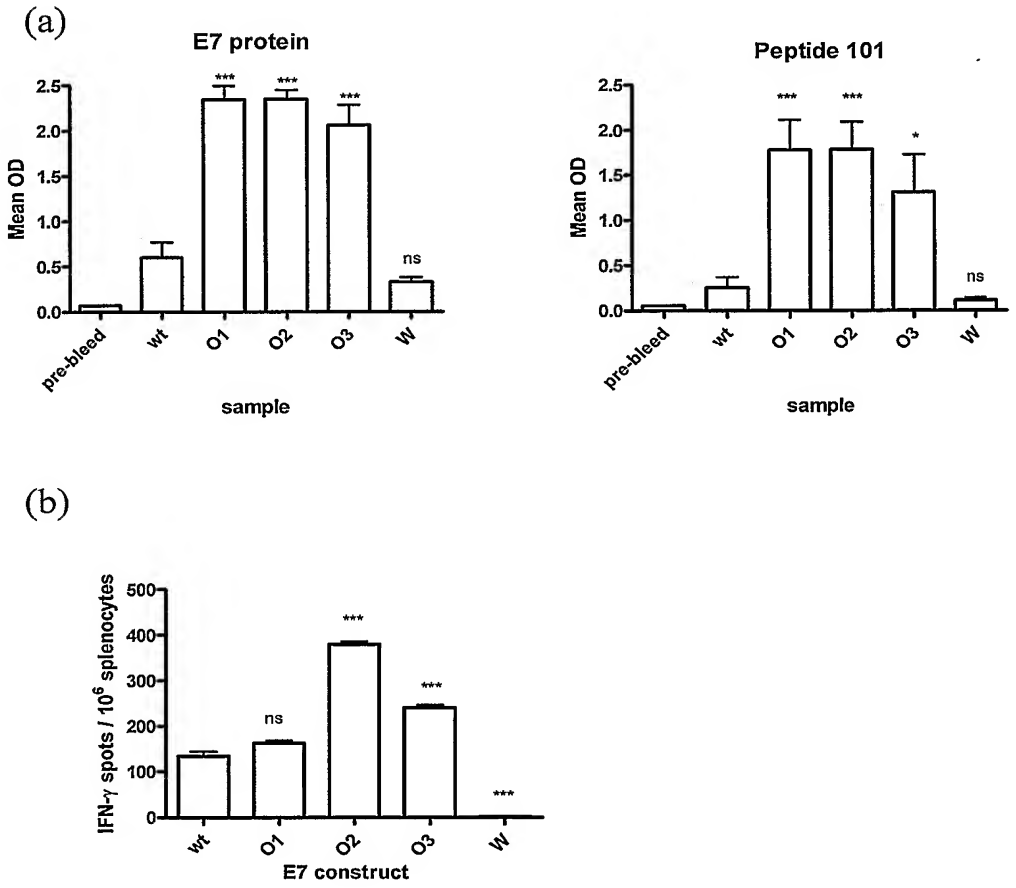
(a)
E7 protein
Mean OD
2.5
2.0
1.5
1.0
0.5
0.0
***
***
***
ns
pre-bleed
wt
O1
O2
O3
W
sample
Peptide 101
Mean OD
2.5
2.0
1.5
1.0
0.5
0.0
***
***
*
ns
pre-bleed
wt
O1
O2
O3
W
sample
(b)
IFN-γ spots / 10^6 splenocytes
500
400
300
200
100
0
ns
***
***
***
wt
O1
O2
O3
W
E7 construct

FIGURE 18

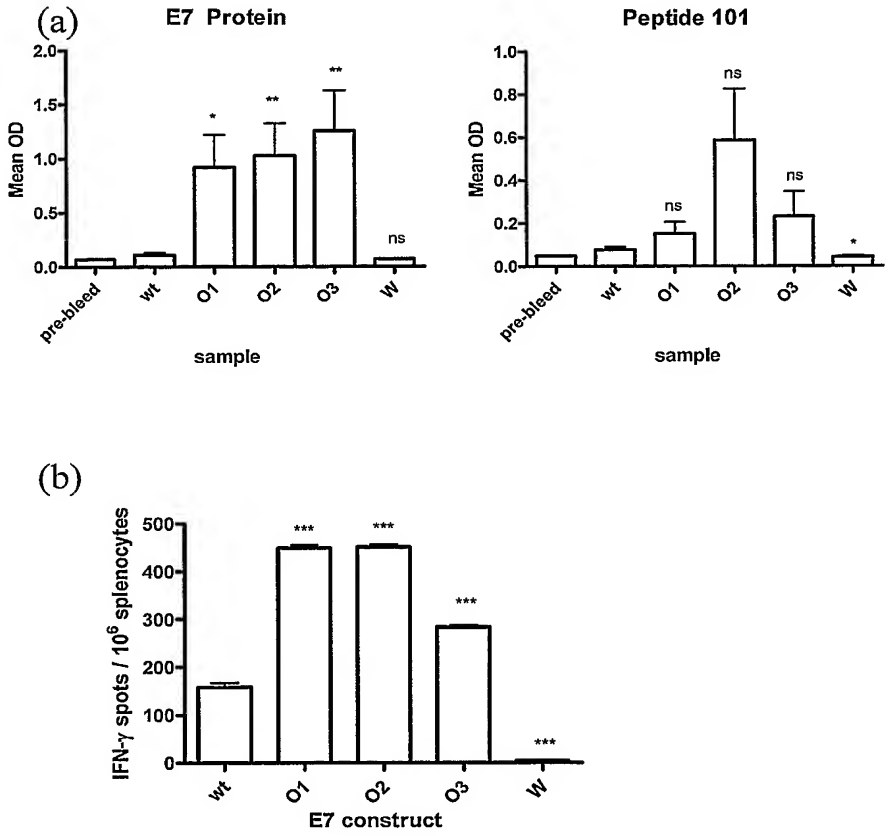
(a)
E7 Protein
Peptide 101
Mean OD
2.0
1.5
1.0
0.5
0.0
1.0
0.8
0.6
0.4
0.2
0.0
*
**
**
ns
ns
ns
*
pre-bleed
wt
O1
O2
O3
W
sample
(b)
IFN-γ spots / 10^6 splenocytes
500
400
300
200
100
0
***
***
***
***
wt
O1
O2
O3
W
E7 construct

FIGURE 19

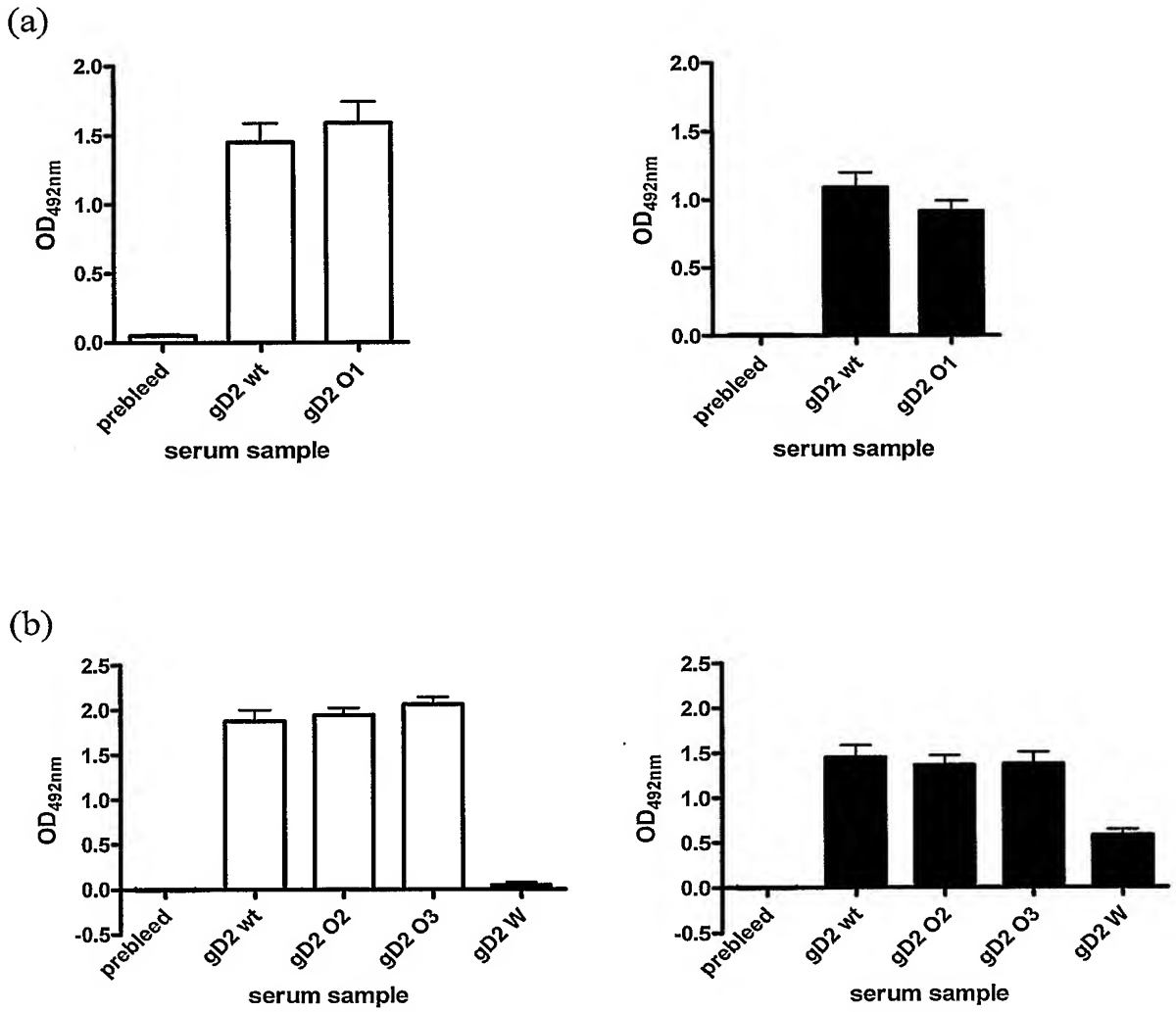
(a)
2.0
1.5
1.0
0.5
0.0
OD492nm
prebleed
gD2 wt
gD2 O1
serum sample
2.0
1.5
1.0
0.5
0.0
OD492nm
prebleed
gD2 wt
gD2 O1
serum sample
(b)
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
OD492nm
prebleed
gD2 wt
gD2 O2
gD2 O3
gD2 W
serum sample
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
OD492nm
prebleed
gD2 wt
gD2 O2
gD2 O3
gD2 W
serum sample

FIGURE 20